

**¿EXISTEN *PEER EFFECTS* EN EL RENDIMIENTO ACADÉMICO DE LOS  
ESTUDIANTES DE GRADO 11 EN LOS COLEGIOS OFICIALES DE LA  
CIUDAD DE MEDELLÍN?**

**SANDRA MILENA CHICA GÓMEZ**

Trabajo de grado para optar el título de Magister en Economía

Asesor

**Mónica Patricia Ospina Londoño.**

**Universidad EAFIT  
Escuela de Administración  
Departamento de Economía  
Medellín  
2011**

## Contenido

Introducción .....	3
1. Revisión de la literatura. ....	6
2. Presentación de la hipótesis.....	9
3. Modelo conceptual. ....	10
4. Metodología. ....	11
4.1. MCO y regresión por cuantiles.....	12
4.2. Modelo Lineal Jerárquico (HLM). ....	14
5. Datos .....	16
6. Resultados .....	23
6.1. Matemáticas MCO y cuantiles.....	23
6.2. Lenguaje MCO y cuantiles.....	29
6.3. Modelos lineales jerárquicos (HLM).....	33
7. Conclusiones .....	35
Bibliografía .....	38
Anexos .....	43

## Contenido de tablas

Tabla 1. Evolución tasa de cobertura escolar en Medellín. (Educación media).....	4
Tabla 2. Diccionario de variables.....	18
Tabla 3. Edad .....	19
Tabla 4. Computador .....	20
Tabla 5. DVD .....	20
Tabla 6. Servicio de TV .....	20
Tabla 7. Internet .....	20
Tabla 8. Nivel educación padre .....	21
Tabla 9. Proviene de colegio privado .....	22
Tabla 10. Personas en el hogar.....	22
Tabla 11. Resultado de matemáticas y lenguaje .....	22
Tabla 12. Estadísticas descriptivas de las variables dependientes.....	23
Tabla 13. Estimación por MCO para el área de matemáticas. ....	26
Tabla 14. Estimación por MCO para el área de lenguaje.....	30
Tabla 15. Modelo Jerárquico lineal para matemáticas y lenguaje .....	33
Tabla 16. Test de Homocedasticidad matemáticas. ....	43
Tabla 17. Test de Homocedasticidad lenguaje. ....	43
Tabla 18. Varianza Inflada (VIF) matemáticas .....	43
Tabla 19. Varianza Inflada (VIF) lenguaje .....	44
Tabla 20. Test Wald matemáticas .....	44
Tabla 21. Test Wald lenguaje .....	44
Tabla 22. Estimación por cuantiles para el área de matemáticas .....	44
Tabla 23. Estimación por cuantiles para el área de lenguaje .....	46

# **¿Existen *peer effects* en el rendimiento académico de los estudiantes de grado 11 en los colegios oficiales de la ciudad de Medellín?**

**Sandra M. Chica G<sup>1</sup>**

## **Introducción**

La preocupación por la educación ha dejado de centrarse en la cobertura para trascender al campo de la calidad educativa. Es así, como numerosos estudios han centrado su análisis en los resultados arrojados por diferentes pruebas estándar como mecanismo para comprender los determinantes del desempeño académico de los estudiantes. La discusión sobre dichos determinantes ha sido amplia, se han identificado factores asociados a la escuela, a la calidad de los profesores, al gasto en educación y a aspectos socioeconómicos de los estudiantes (Coleman et al (1966). Jenks (1972) Alexander y Simmons (1975), Toutkoushian y Curtis (2005), Wolff Schiefelbein y Valenzuela (1993). Sin embargo, investigaciones internacionales han incluido, adicional a las variables tradicionales de factores asociados a la institución educativa y la situación socioeconómica de los estudiantes, un nuevo elemento que podría llegar a complementar el análisis sobre las variables que influyen en el rendimiento académico de los estudiantes y permitir una mayor comprensión del problema. Este elemento es conocido en la literatura como *peer effects*.

Los efectos de pares o *peer effects* es un término que se deriva de lo que comúnmente se conoce en la literatura como interacciones sociales<sup>2</sup>. Éstos pueden dividirse en tres tipos: el primero, llamado *efecto endógeno*, se refiere a que las acciones de un individuo se ven afectadas por el comportamiento o acciones de sus compañeros; para el caso, el rendimiento de un estudiante depende del rendimiento promedio de sus compañeros de curso. El segundo, llamado *efecto de contexto*, es entendido como la influencia que tiene sobre el

---

<sup>1</sup>Ingeniera Administradora, Universidad Nacional (sede Medellín). Estudiante de Maestría en Economía de la Universidad EAFIT. Correo electrónico: [smchica@gmail.com](mailto:smchica@gmail.com)

<sup>2</sup> Las interacciones sociales se refiere al estudio de los mecanismos por los cuales la sociedad afecta al individuo (Manski, 1995). Si se define la sociedad como "los compañeros de aula", las interacciones sociales toman el nombre de *peer effects*.

desempeño de un estudiante las características de sus compañeros de clase (por ejemplo edad, sexo y características de los padres), y el tercero llamado *efecto correlacionado*, donde los individuos en un grupo se comportan de manera similar porque enfrentan un ambiente similar<sup>3</sup>.

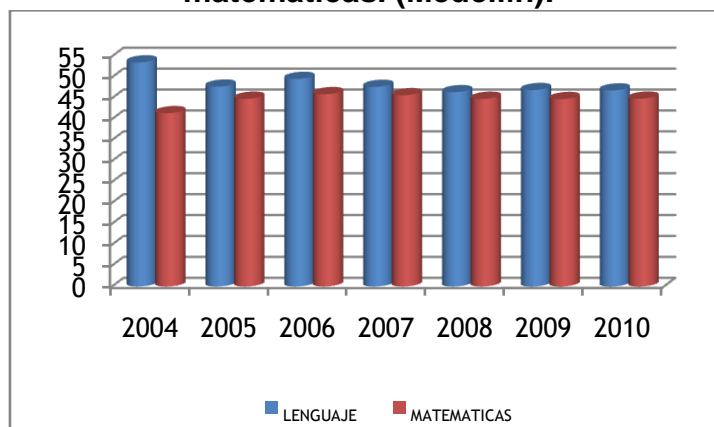
La ciudad de Medellín desde hace un poco más de siete años ha sufrido grandes transformaciones, y una de las más significativas ha girado en torno a la preocupación por el mejoramiento de la calidad educativa bajo el lema *Medellín La Más educada*. Sin embargo, a pesar de que en cobertura se ha logrado avanzar (Ver Tabla 1), los resultados en las pruebas nacionales no han tenido grandes variaciones frente a los logros alcanzados en materia de calidad, como se puede observar en el Gráfico 1, los resultados promedio para esta ciudad han mejorado levemente en lo relacionado con el área de matemáticas, pero disminuido en lenguaje en los últimos seis años.

**Tabla 1. Evolución tasa de cobertura escolar en Medellín. (Educación media).**

	2004	2005	2006	2007	2008	2009	2010
<b>Tasa de cobertura neta<sup>4</sup></b>	45,2%	51,9%	52,7%	52,5%	53,3%	56,3%	58,2%
<b>Tasa de cobertura bruta</b>	61,4%	67,1%	67,3%	71,0%	72,6%	76,2%	79,9%

**Fuente:** Secretaría de Educación; Subsecretaría de Planeación; Sistema de Matrícula en línea.

**Gráfico 1. Evolución del resultado promedio, área de lenguaje y matemáticas. (Medellín).**



**Fuente:** ICFES Interactivo.

<sup>3</sup>Manski, (1995).

<sup>4</sup>La tasa de cobertura neta es el resultado de restar los estudiantes matriculados que se encuentran en extraedad.

En este sentido, el objetivo de este estudio es determinar la existencia y magnitud de *peer effects* en las instituciones educativas oficiales de la ciudad de Medellín, tomando como proxy del rendimiento académico los resultados de la prueba Saber 11° para el segundo semestre de 2010. Para esto se utiliza una base de datos resultado de cruzar los formularios de registro y resultado de las pruebas Saber 11°, y el Sistema de Matrícula en línea de la ciudad de Medellín, para ese año. Lo anterior, se justifica porque esta última base permite identificar dentro del colegio, el grupo en el cual se encuentra cada estudiante. El análisis realizado se centra en los resultados de las áreas de matemáticas y lenguaje, ya que estas áreas son consideradas como fundamentales en el aprendizaje.

La importancia de este trabajo radica en que no existe evidencia de un trabajo que haya estudiado los *peer effects* en Colombia ni en la ciudad de Medellín. El estudio de esta variable puede generar importantes herramientas de políticas encaminadas a potencializar la conformación de las aulas de clase de manera que la redistribución de los grupos, teniendo en cuenta las características de los estudiantes, vaya en pro de mejorar el rendimiento. Estas estrategias tienen la ventaja de que no requerirían grandes inversiones como lo hacen otro tipo de políticas, además pueden ser puestas en marchas por cada rector o institución lo cual minimiza en gran medida los trámites que se pueden generar.

Para alcanzar el objetivo se realizan tres modelos, el primero mediante una estimación por mínimos cuadrados ordinarios(MCO), el cual permite identificar la existencia de *peer effects* y otras variables explicativas que influyen el rendimiento académico, un segundo ejercicio a partir de la metodología de regresiones por cuantiles y un tercer análisis usando las metodologías de modelos jerárquicos (HLM). La regresión por cuantiles permite explorar cuáles variables afectan más a los estudiantes con puntajes bajos y cuáles influyen en los de puntajes altos, el cual sirve como herramienta para focalizar políticas, teniendo en cuenta el nivel de logro alcanzado por los estudiantes. Los dos primeros modelos permitieron identificar la existencia de efectos de pares endógenos en los colegios oficiales de la ciudad, es decir la influencia que tiene

el tener compañeros de clase con buenos desempeños, tanto para el rendimiento en matemáticas como en lenguaje. La estimación por Modelos Jerárquicos Lineales (HLM) permite conocer el coeficiente de correlación intraclase y determinar qué tan necesario es que futuras investigaciones en Colombia tengan en cuenta la estructura jerárquica de los procesos educativos y aborden el tema de efectos de pares a través de la estimación de HLM.

Este trabajo se encuentra dividido en siete secciones adicional a ésta. En la primera se realiza una revisión de la literatura sobre los *peer effects* y otras variables que influyen en el rendimiento académico de los estudiantes y en la segunda se presenta la hipótesis sobre la cual parte este trabajo. En la tercera sección se muestra una descripción del modelo conceptual, el cual toma como punto de partida la determinación de una función de producción de la educación. En la cuarta se exponen los modelos econométricos a estimar, en la quinta se describen los datos utilizados y en la sexta se presentan los principales resultados. Por último, la séptima sección da a conocer las principales conclusiones y recomendaciones del estudio.

## **1. Revisión de la literatura.**

Han sido varias las ramas del conocimiento que se han preocupado por la importancia que tienen las interacciones sociales en el comportamiento y desempeño de los individuos. La psicología, sociología y más recientemente la economía, han incluido en sus estudios la influencia que tienen las características de los vecinos o compañeros de grupos en el desempeño. En el caso particular del desempeño académico se han desarrollado diferentes investigaciones, sin embargo aún no se ha llegado a un consenso en cuanto a qué tan importante es esta interacción o efectos de pares en el desempeño académico de los estudiantes.

En Dinamarca, Schindler (2000) combina los datos de PISA 2000 con los registros del Sistema del país con el fin de estudiar los efectos de pares. En este trabajo se realiza una regresión por Mínimos Cuadrados Ordinarios (MCO) y luego por cuantiles para determinar cómo se comportan los efectos de pares

a lo largo de la distribución del resultado. La autora encuentra evidencia a favor de los efectos de pares en el país, así un aumento en el logro de educación promedio de los pares es asociado con un incremento de la puntuación en lectura en casi 5 puntos en la escala PISA, lo que equivale a un 5% de la desviación estándar de la escala, resultado muy similar al que se encuentra en este trabajo.

En Hattie (2002), se estudian los efectos de pares teniendo en cuenta tres posibilidades de organización del salón de clase: se basa en la capacidad, en la etnia o en el género, e incluye además los efectos de las clases multigrado, de distintas edades y tamaño de clase. A pesar de que el autor encuentra evidencia a favor de la existencia de efectos de pares, advierte que la enseñanza de calidad es más influyente y parece ser independiente de la configuración de la clase o la homogeneidad de los estudiantes dentro de la misma.

McEwan (2003), presenta un trabajo en el cual estima el efecto de pares en Chile para estudiantes de octavo grado de varios salones de clase en una misma escuela. Con el fin de minimizar el riesgo de endogeneidad de las variables, realiza una regresión teniendo en cuenta efectos fijos, controlando los efectos no observados mediante variables familiares y propias del estudiante. Concluye que el efecto de pares es significativo y la variable que más influye es la educación media las madres, así un aumento de una desviación estándar conduce a una ganancia en 0.27 desviaciones estándar, sin embargo esta educación está sujeta a rendimientos marginales decrecientes.

Robertson D. & Symons J (2003), analizan en este artículo la existencia de *peer effects* utilizando la encuesta del Estudio Nacional de Desarrollo Infantil en el Reino Unido, mediante la aplicación de un modelo Tobit con los resultados en las áreas de matemáticas y competencia lectora de estudiantes de 11 años. Los autores tienen en cuenta en su análisis, variables relacionadas con la educación, ingresos y clase social de los padres. En el estudio se encuentra fuerte evidencia a favor de la importancia de los grupos de pares y más aún

cuando se tiene en cuenta la clase social y los logros académicos de los padres de los compañeros de clase.

Hanushek, E.A., Kain, J.F., Markman, J.M., Rivkin, S.G., (2003), analizan si los pares tienen la capacidad de afectar el rendimiento académico de los estudiantes. Para ello utilizan un panel de datos con estudiantes de tres cohortes sucesivas de escuelas primarias públicas de Texas, comenzando con los estudiantes que asistieron a tercer grado en 1992. Utilizan el método de ecuaciones simultáneas bajo el marco de efectos fijos para realizar las estimaciones. El mayor hallazgo está en encontrar que el desempeño promedio de los pares tiene influencia en el aprendizaje de los estudiantes. Un aumento de 0,1 desviaciones estándar en el rendimiento medio de los pares conduce a un aumento de aproximadamente 0,02 en el logro.

Lefgren (2004), realiza un estudio sobre en qué medida la calidad de compañeros de clase es determinante del rendimiento académico de un estudiante (*peer effects*). Para ello utiliza la metodología de variables instrumentales para una base de datos de 150.000 estudiantes de tercero y sexto grado en Chicago. En este estudio se tiene en cuenta que en Estados Unidos se pueden presentar dos tipos de escuela: con seguimiento y sin seguimiento. En las escuelas sin seguimiento, se encuentran estudiantes con capacidades similares en el mismo salón, mientras que en las escuelas con seguimiento, los estudiantes son asignados a las aulas con base en sus logros académicos. El análisis que realiza arroja como resultado que aunque existe evidencia de que existen efectos de pares, estos son bastante modestos y no terminan siendo tan importantes a la hora de acciones de política pública.

Bonesrønning (2006), analiza un caso particular de los *peer effects*, en el cual estudia el efecto que tiene sobre el rendimiento académico el hecho de que un estudiante comparta su salón de clases con compañeros pertenecientes a familias disueltas. Para ello utiliza los datos de estudiantes de secundaria en Noruega y realiza una regresión lineal por MCO. Los resultados arrojan evidencia a favor de que el desempeño de los estudiantes de secundaria se ve



afectado negativamente por la presencia de compañeros de clase con familias disueltas.

Para el caso colombiano, si bien se han realizado diversos estudios sobre determinantes del rendimiento académico, donde sólo se estudian variables socioeconómicas, personales e institucionales: Gaviria y Barrientos (2001a), (2001b) y (2008), Correa (2004), Cepeda (2005) y Tobón et al. (2008); ninguno de ellos ha tenido en cuenta el efecto que podrían tener las características y desempeño de los pares en los logros alcanzados. Asimismo, Medellín ha sido una de las ciudades que más ha invertido en educación en los últimos años, y sería interesante poder conocer qué tanto podría influir en los colegios oficiales determinar la distribución de los salones de clase teniendo en cuenta las características y desempeño promedio de los estudiantes, algo que se utiliza en otros países, pero que ni en Colombia ni en la ciudad se ha tenido en cuenta.

## **2. Presentación de la hipótesis.**

**Hipótesis:** Existen efectos de pares en el rendimiento académico de los estudiantes de grado 11° en las instituciones educativas oficiales de la ciudad de Medellín.

Teniendo en cuenta que el estudio de las interacciones sociales ha tomado gran fuerza en los últimos años y en la mayoría de los casos ha arrojado evidencia a favor del hecho de que el comportamiento y las características de los individuos que nos rodean ejercen un efecto sobre el desempeño o comportamiento de cada uno, este trabajo tiene como hipótesis que lo anterior también se cumple en el ámbito educativo en la ciudad de Medellín.

La literatura ha demostrado con diferentes estudios la existencia de factores asociados al desempeño académico (factores socioeconómicos, personales, académicos y de escuela), en tal sentido, partiendo de la idea de que en distintos contextos las interacciones sociales se han evidenciado como determinantes a la hora de definir el comportamiento y actitudes de los

individuos, se pretende establecer cómo las características de los compañeros de clase juega un papel determinante en el desempeño.

### **3. Modelo conceptual.**

La teoría económica ha tenido grandes avances, el concepto de crecimiento y desarrollo económico ha evolucionado en los últimos años. El modelo tradicional de crecimiento de Solow (1956) plantea el crecimiento económico en términos de los factores tradicionales de producción (capital físico y trabajo) y la tecnología como variable exógena, el cual fue pieza clave para el desarrollo de nuevas teorías. El punto de partida de los nuevos estudios se basaba en la limitación que tenía el modelo de Solow, al dejar una buena parte del crecimiento que no era explicado por estos dos factores, lo que comúnmente se conoce en la literatura como el residuo de Solow. Es aquí donde comienzan a tomar fuerza las teorías que incluyen el capital humano como otro factor determinante del crecimiento.

Uno de los primeros exponentes del concepto de capital humano fue Schultz en su obra “Investment in Human Capital” (1961), sin embargo es con el trabajo de Becker (1964) donde toma mayor fuerza. Becker describe el capital humano como las habilidades y destrezas que las personas van adquiriendo a lo largo de su vida, bien sea por medio de estudios formales, como las escuelas, o por conocimientos informales, que son lo que adquieren las personas por medio de la experiencia. La introducción del concepto de capital humano como factor determinante del crecimiento económico aparece entonces en el modelo planteado por Mankiw, Romer y Weil (1992), quienes complementan el modelo de Solow con este nuevo factor. En este sentido, la educación como medida de capital humano ha motivado diversos estudios económicos, a tal punto que se ha establecido todo un desarrollo teórico frente al desarrollo de la “Función de Producción Educativa (FPE)”.

En la teoría económica la función de producción establece la relación existente entre los factores o insumos utilizados en un proceso productivo (inputs), y el producto obtenido (outputs), dada cierta tecnología. La analogía con la FPE

establece que el proceso educativo puede asimilarse al proceso productivo de cualquier bien o servicio: existen factores e insumos que, combinados de diferentes maneras, dan lugar a diferentes cantidades y calidades de bien final o producto terminado. La preocupación en la mayoría de estudios sobre factores asociados a la educación gira en torno al producto final en términos de calidad o rendimiento académico, es decir, de qué manera pueden combinarse los insumos o entradas existentes para obtener mayores rendimientos de los estudiantes. Partiendo de este concepto muchos autores han enfocado sus estudios sobre rendimiento académico, o retornos de la educación en la estimación de la FPE, se destacan los trabajos realizados por Hanushek (1979) y (1989), Hedges (1994), Henderson (1978), Lazear (2001), entre otros.

Este trabajo parte entonces de este marco conceptual, para lo cual se estimará una función de producción que recoja las variables a tener en cuenta, representada de la siguiente manera:

$$\textbf{Rendimiento} = f(\textit{caract. Individuales, caract, socioeconómicas, variables de colegio, características promedio de los pares})$$

$$y_{icj} = f(X_{icj}, S_{icj}, R_{icj}, P_{cj})$$

Teniendo en cuenta lo anterior, el rendimiento académico del estudiante  $i$  en el grupo del colegio  $J$ , medido como resultado de las pruebas Saber 11º ( $y_{ij}$ ), estará en función de las características individuales ( $X_i$ ), socioeconómicas ( $S_i$ ), del colegio (recursos) ( $R_i$ ), y características promedio de los pares (*peer effects*) ( $P_i$ ).

#### 4. Metodología.

Para corroborar la hipótesis de este trabajo, se utilizaron tres metodologías. La primera utilizando el modelo de regresión lineal por Mínimos Cuadrados Ordinarios (MCO) para tratar de encontrar las variables que impactan de manera significativa el desempeño, luego se realiza la regresión por cuantiles para el modelo de MCO con la cual se pretende encontrar las diferencias de los

efectos de las variables entre los diferentes cuantiles de la distribución. Por último se corrió un modelo lineal jerárquico (HLM) de dos niveles, con el fin de dejar abierta la posibilidad de estudiar los efectos de pares teniendo en cuenta la estructura jerárquica de los datos del estudio.

#### **4.1. MCO y regresión por cuantiles.**

El método de MCO es muy utilizado en la literatura económica y tiene como objetivo explicar el comportamiento de una variable (endógena) utilizando los valores tomados por un conjunto de variables explicativas. Por su parte, el método de regresión por cuantiles<sup>5</sup> intenta modelar el efecto que tienen las variables explicativas sobre toda la distribución condicional de  $y$ , a diferencia de los modelos de regresión estándar como MCO, el cual se concentra sólo en la esperanza condicional. Esta metodología relaja los supuestos del modelo anterior (MCO) respecto al comportamiento de los errores y estima los efectos de las variables independientes sobre la mediana en distintos puntos de la distribución condicional de la variable dependiente.

Koenker y Bassett (1978), en un trabajo de crucial importancia, proponen el siguiente modelo de regresión para la distribución condicional de:

$$Q_{y|x}(t) = x'\beta(\tau)$$

En donde  $\tau \in (0,1)$  y  $\beta(\tau)$  es un vector de  $K$  coeficientes. La notación  $Q_{y|x}(\tau)$  hace referencia al  $(\tau)$ -ésimo cuantil de la distribución de  $y$  condicional en  $x$ . Es importante resaltar que para una variable aleatoria  $Z$  con función de distribución acumulada  $F(z)$  continua y monótona,  $(\tau)$ -ésimo cuantil es un número  $Q_Z(\tau)$  que satisface:

$$F(Q_Z(\tau)) = \tau$$

Para llevar a cabo el estudio se hizo la siguiente especificación del modelo a estimar, la cual se realizará tanto para el resultado en matemáticas como en lenguaje.

---

<sup>5</sup> Sosa W. (2005) En: Progresos en econometría. Págs. 101-108.

$$\begin{aligned}
\text{Resultado}_{icj} = & \beta_{0,j} + \beta_{1,icj} \text{ genero} + \beta_{2,icj} \text{ edad} + \beta_{3,icj} \text{ trabaja} + \beta_{4,icj} \text{ tienecomputador} \\
& + \beta_{5,icj} \text{ reproductorDVD} + \beta_{6,icj} \text{ servicioTV} + \beta_{7,icj} \text{ internet} \\
& + \sum_{n=0}^5 \beta_{8+n,icj} \text{ ingresos} + \sum_{l=0}^8 \beta_{14+l,icj} \text{ nivelpadre} + \sum_{h=0}^8 \beta_{23+h,icj} \text{ nivelmadre} \\
& + \beta_{31,icj} \text{ provieneprivado} + \beta_{32,icj} \text{ personashogar} + \sum_{s=0}^3 \beta_{33+s,icj} \text{ jornada} \\
& + \beta_{37,icj} \text{ prommatematicas} + \beta_{38,icj} \text{ promlenguaje}
\end{aligned}$$

Una de las principales limitaciones que se tiene al modelar *peer effects* por medio de MCO es el sesgo de autoselección en el cual se puede incurrir, ya que las familias por lo general toman la decisión de matricular a sus hijos basados en ciertos criterios como el desempeño general del plantel. Sin embargo, en el caso de los colegios oficiales de la ciudad de Medellín, los padres escogen la institución en la mayoría de los casos basándose en criterios de cercanía al colegio, es por esto que este sesgo se trata de controlar por medio de la inclusión de variables relacionadas con las características socioeconómicas y familiares de los estudiantes y algunas relacionadas con la institución educativa. Por otra parte, al interior de las aulas y la organización de los salones de clase también se corre el riesgo de tener problemas de selección, en algunos países existen criterios de agrupación interna en los salones de clase, sin embargo para el caso de colegios oficiales de la ciudad de Medellín no existen criterios establecidos y en la mayoría de los casos los estudiantes son incluidos en las aulas de clase aleatoriamente de manera que lo que se busca es ir cubriendo los cupos disponibles.

Otra limitación que se tiene al estimar *peer effects* es el problema de endogeneidad que se genera con las variables promedio, ya que el alumno *i* se ve afectado por el promedio de clase, y este alumno es a su vez parte del promedio de esta clase, impactando de esta manera a los demás alumnos. Este problema se corrige en el trabajo sacando del promedio del grupo, el resultado del alumno *i*, lo anterior se considera una ventaja en este estudio, ya que se elimina el sesgo de endogeneidad en el resultado.

Un último aspecto que se debe aclarar es que para las bases de datos disponibles no se cuenta con información que permita controlar por el efecto “maestro” lo cual podría generar inquietudes frente a un posible sesgo por variable omitida, ya que el efecto de pares podría estar sobreestimado al no contarse con este control. Sin embargo, al estudiarse un poco la organización y funcionamiento del sistema educativo en Medellín específicamente para las instituciones oficiales, se encuentra que en la mayoría de las instituciones el profesor de matemáticas y lenguaje asignado es el mismo para los diferentes grupos del plantel educativo. Por ejemplo, dado que se tienen horarios designados por área para cada grupo, en su mayoría un mismo profesor de matemáticas es el encargado de dictar las diferentes clases para los grupos de grado 11º. Lo anterior permite corroborar que en este caso el efecto maestro no estaría sobrevalorando los *peer effects* encontrados, ya que el aporte que el maestro pueda hacer al rendimiento es el mismo para los diferentes grupos de un mismo colegio. Lo que si ocurre es encontrar diferencias entre los aportes de los maestros en los diferentes colegios debido a que son diferentes docentes los que imparten las clases, sin embargo para este estudio se cuentan con variables que permiten controlar por las características institucionales.

#### **4.2. Modelo Lineal Jerárquico (HLM).**

La Modelación Lineal Jerárquica es una técnica de regresión la cual toma en cuenta la estructura jerárquica de los datos, para el caso de este trabajo se asume que los estudiantes están anidados en grupos de una misma escuela. En esta situación, es razonable postular un modelo que considere una posible diferencia entre las unidades de nivel 2, es decir, plantear un modelo tal que, para cada unidad en el nivel 2 se tengan diferentes coeficientes. Bajo esta situación el modelo lineal jerárquico permite simultáneamente hacer un estudio en los niveles de la estructura jerárquica, tomando en cuenta variables para las unidades en cada uno de los niveles. En los modelos lineales jerárquicos cada

uno de los niveles de la estructura jerárquica es representado formalmente con su propio submodelo<sup>6</sup>.

Para este estudio se partirá del modelo nulo el cual es totalmente incondicional y no tiene en cuenta variables explicativas en ninguno de los niveles y luego se realizará un segundo modelo con una variable explicativa a nivel 1 con el fin de conocer el coeficiente de correlación intraclase y determinar la posibilidad de tener en cuenta este tipo de modelos en futuros estudios para identificar el efecto de pares en Colombia.

**Modelo 1:** nulo o incondicional.

$$\begin{aligned} \text{Nivel 1:} & \quad \text{Resultado}_{ij} = \beta_{0j} + e_{ij} \\ \text{Nivel 2} & \quad \beta_{0j} = \gamma_{00} + U_{0j} \\ \text{Modelo mixto:} & \quad \text{Resultado}_{ij} = \gamma_{00} + U_{0j} + e_{ij} \end{aligned}$$

$$\rho = \frac{\tau_{00}}{\tau_{00} + \sigma^2}$$

**$\rho$  = coeficiente de correlación intraclase (CCI)**

$\sigma^2$  = es la varianza al interior de los grupos de los colegios.

$\tau_{00}$  = es la varianza entre los grupos de los colegios.

Si  $\rho = 0$ , colegios o grupos homogéneos.

Si  $\rho = 1$ , no hay variabilidad al interior de los grupos, estudiantes iguales.

**Modelo 2:** con variable explicativa a nivel, pero se asume con efectos fijos a nivel dos.

$$\begin{aligned} \text{Nivel 1:} & \quad \text{Resultado}_{ij} = \beta_{0j} + \beta_{1j} \text{genero}_{1ij} + e_{ij} \\ \text{Nivel 2:} & \quad \beta_{0j} = \gamma_{00} + U_{0j} \\ & \quad \beta_{1j} = \gamma_{10} \end{aligned}$$

$$\text{Modelo mixto: } \text{Resultado}_{ij} = \gamma_{00} + \gamma_{10} \text{genero}_{1ij} + U_{0j} + e_{ij}$$

---

<sup>6</sup> Smith, (2011).

En los modelos HLM el análisis se realiza partiendo de la matriz de varianzas y covarianzas  $\tau$ :

$$\tau = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix}$$

La matriz  $\tau$  es una matriz simétrica, la diagonal representa las varianzas y los otros elementos representan las covarianzas entre las pendientes y los interceptos. Cuando se asume que los interceptos y las pendientes son mutuamente independientes se fija  $\tau_{10} = \tau_{01} = 0$ .

Los modelos HLM presentan ventajas sobre las estimaciones con MCO, entre ellas se encuentra que permiten estudiar las relaciones entre variables teniendo en cuenta la inter-dependencia entre observaciones pertenecientes a un mismo grupo. Genera además estimaciones más precisas, y evita caer en la paradoja de Simpson la cual se da cuando ocurre la desaparición de una asociación o comparación significativa de dos variables cuando los datos son desagregados por grupos. Al modelar HLM se deben tener en cuenta varias estimaciones incluyendo o no variables dependientes en algunos de los niveles, la elección del mejor modelo de estimación se basará en el menor valor del Criterio de Información de Akaike (AIC) o el menor valor del Criterio de información Bayesiano de Schwarz<sup>7</sup>, estos son los criterios tenidos en cuenta en este trabajo.

## 5. Datos

Los datos utilizados en este estudio son el resultado del cruce de dos bases de datos: la base de datos correspondiente a los resultados y registro de las pruebas de estado ICFES Saber 11<sup>o</sup> para el segundo semestre de 2010 y la información recopilada en el Sistema de Matrícula en Línea de la Secretaría de Educación de Medellín (SEM), correspondiente a la Matrícula oficial y contratada validada por el Ministerio de Educación Nacional, anexo 6A con fecha de corte a noviembre 2 de 2010. La variable de interés del presente

---

<sup>7</sup> Martínez (2011).



estudio será el resultado de la evaluación en las principales áreas del conocimiento: matemáticas y lenguaje.

La base de datos correspondiente a la aplicación de la Prueba Saber 11 para el segundo semestre de 2010 contiene en total 31686 registros de estudiantes<sup>8</sup> para la ciudad de Medellín incluyendo colegios privados y oficiales; y la del Sistema de Matrícula en Línea de la SEM contiene 22287 registros para grado 11 de la misma ciudad. Dado que los datos del ICFES son para todos los estudiantes que hayan presentado la prueba en la ciudad de Medellín y el Sistema de Matrícula en Línea sólo contiene datos de los colegios oficiales y contratados, los registros efectivos al calcular el cruce fueron de 17733, logrando un porcentaje de coincidencia de aproximadamente el 80%. No obstante, para las variables del nivel educativo de los padres se realizó una imputación para 2139 datos<sup>9</sup> (1671 para el padre y 468 para la madre) con el fin de no perder información importante de la base de datos, lo cual garantiza un buen tamaño de muestra y confiabilidad de los resultados de los modelos estimados. La imputación se realizó teniendo en cuenta el mismo colegio, grupo y nivel de ingresos.

Los datos de registro de la Prueba Saber ICFES de 2010 se pueden agrupar en tres categorías: datos personales, información académica (variables propias de la institución educativa) e información familiar (variables de contexto). Adicional a esto contiene los datos de resultado de los estudiantes para las diferentes áreas evaluadas. A pesar de que el ICFES y el Sistema de Matrícula en Línea reportan algunas variables comunes, de esta última sólo se tomaron las variables: código DANE de la institución, nombre del estudiante, identificación y tipo de identificación; variables que permitieron el cruce de información entre las bases de datos. Adicional a estas, se tuvo en cuenta si el estudiante proviene del sector privado y la variable que define el grupo de cada estudiante, la cual permite estudiar la hipótesis de esta investigación. La inclinación por las variables del ICFES está fundamentada en que éste es un

---

<sup>8</sup>Este número se tiene en cuenta omitiendo aquellos estudiantes que ya han presentado la prueba Saber 11<sup>o</sup>

<sup>9</sup> La imputación se realizó teniendo en cuenta el promedio del nivel educativo de los padres por colegio, grupo de ingresos mensuales del estudiante.

registro más reciente y por lo tanto puede contener información personal y familiar más actualizada del estado del estudiante, teniendo en cuenta que en la mayoría de casos en el Sistema de Matrícula en Línea sólo se renuevan datos como grado, grupo e institución.

Las variables utilizadas en la estimación se exponen en la Tabla 2. En la última columna se señala, para el caso de las variables categóricas que toman más de dos valores, aquella que se escoge como variable omitida.

**Tabla 2. Diccionario de variables**

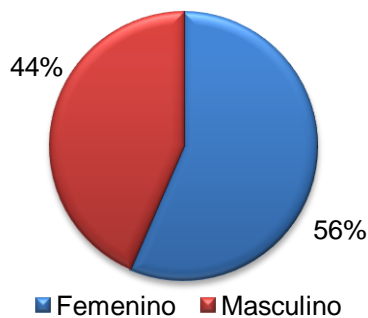
Variable		Descripción de la variable	Variable omitida
<b>NM</b>		Nivel matemáticas	
<b>NL</b>		Nivel lenguaje	
<b>Genero</b>	0	Femenino	
	1	Masculino	
<b>Edad</b>		Años	
<b>Trabaja</b>	0	No	
	1	Si	
<b>Área residencia</b>	0	Rural	
	1	Cabecera	
<b>Tiene computador</b>	0	No	
	1	Si	
<b>Reproductor de DVD</b>	0	No	
	1	Si	
<b>Servicio de TV</b>	0	No	
	1	Si	
<b>Internet</b>	0	No	
	1	Si	
<b>Ingresos mensuales</b>	1	Menos de 1 SM	X
	2	Entre 1 y menos de 2 SM	
	3	Entre 2 y menos de 3 SM	
	4	Entre 3 y menos de 5 SM	
	5	Entre 5 y menos de 7 SM	
	6	Entre 7 y menos de 10 SM	
	7	10 o más SM	
<b>Nivel educación padre/Nivel madre</b>	0	Ninguno	X
	1	Primaria Incompleta	
	2	Primaria Completa	
	3	Secundaria (bachillerato) Incompleta	
	4	Secundaria (bachillerato) Completa	
	5	Educación técnica o tecnológica incompleta	
	6	Educación técnica o tecnológica Completa	
	7	Educación Profesional Incompleta	
	8	Educación Profesional Completa	
	9	Postgrado	
<b>Proviene de colegio privado</b>	0	No	
	1	Si	

Personas en el hogar		Número	
Jornada	1	Completa u Ordinaria	X
	2	Mañana	
	3	Noche	
	4	Sabatina-Dominical	
	5	Tarde	
Nivel promedio de desempeño (ICFES) por grupo	Variables de Peer Effects	El nivel promedio de desempeño (ICFES) de todos los estudiantes de cada grupo (matemáticas y lenguaje)	
Nivel promedio educación padres		El nivel promedio de educación de los padres por grupo (promedio para padre y promedio para madre)	
Nivel promedio ingresos familiares		El nivel promedio de ingresos por grupo	

Fuente: Pruebas ICFES Saber 11º y Sistema de Matrícula en Línea.

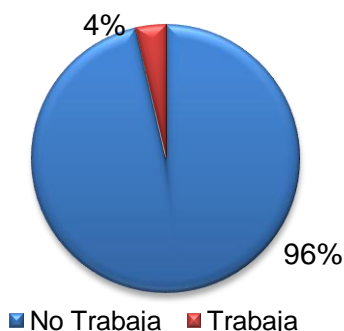
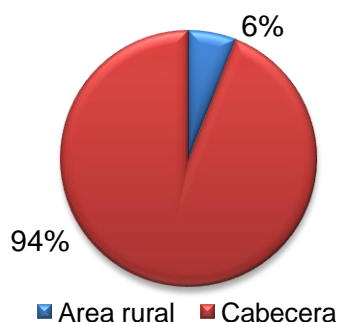
A continuación se muestra la estadística descriptiva de algunas de las variables que se incluirán en el análisis. En el Gráfico 2 se puede observar que el 56% de los estudiantes son mujeres y la edad del total de los estudiantes que presentaron la prueba se concentra entre los 17 y 20 años (Ver Tabla 4). Por otro lado, se observa que la mayoría de los estudiantes no trabaja y que sólo el 6% vive en el área rural, algo que es consecuente con la distribución geográfica de la población en la ciudad de Medellín (Ver Gráficos 3 y 4).

**Gráfico 2. Género**



**Tabla 3. Edad**

Rango edad	Estudiantes	Porcentaje
Menores de 16 años	4424	25%
Entre 17 y 20 años	13177	74%
Mayores de 20 años	132	1%
Total	17733	100%

**Gráfico 3. Trabaja****Gráfico 4. Área residencia**

**Fuente:** Pruebas ICFES Saber 11°, segundo semestre de 2010

En las siguientes tablas se exponen algunos servicios y recursos con los que el estudiante cuenta. Así, el 61% de los estudiantes tienen acceso a un computador y 46% a internet (Ver Tablas 4 y 7); estos dos elementos son considerados, en la mayoría de los casos, como herramientas importantes para facilitar el aprendizaje en los estudiantes. En las Tablas 5 y 6 se exponen que el 66% de los estudiantes tienen DVD y el 72% servicio de televisión por cable, los cuales podrían considerarse como distractores que afectan de manera negativa al tiempo dedicado al estudio.

**Tabla 4. Computador**

	Estudiantes	Porcentaje
<b>No</b>	6974	39%
<b>Si</b>	10759	61%
<b>Total</b>	17733	100%

**Tabla 6. Servicio de TV**

	Estudiantes	Porcentaje
<b>No</b>	5053	28%
<b>Si</b>	12680	72%
<b>Total</b>	17733	100%

**Tabla 5. DVD**

	Estudiantes	Porcentaje
<b>No</b>	5948	34%
<b>Si</b>	11785	66%
<b>Total</b>	17733	100%

**Tabla 7. Internet**

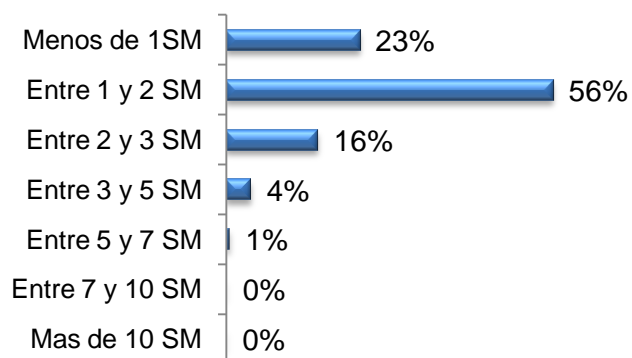
	Estudiantes	Porcentaje
<b>No</b>	9661	54%
<b>Si</b>	8072	46%
<b>Total</b>	17733	100%

**Fuente:** Pruebas ICFES Saber 11°, segundo semestre de 2010

En cuanto a variables socioeconómicas y familiares, se puede observar una estrecha relación entre el nivel educativo alcanzado por los padres y el ingreso

mensual que se percibe en el hogar. Así, aproximadamente el 88% de los padres sólo han culminado sus estudios de secundaria lo cual se ve directamente reflejado en la concentración de ingresos en los niveles más bajos.

**Gráfico 5. Ingresos mensuales**



**Fuente:** Pruebas ICFES Saber 11º, segundo semestre de 2010

**Tabla 8. Nivel educación padre**

Nivel	Padre			Madre		
	Frecuencia	%	Acumulado	Frecuencia	%	Acumulado
Ninguno	699	4%	4%	269	2%	2%
Primaria Incompleta	2847	16%	20%	2423	14%	16%
Primaria Completa	2960	17%	37%	2735	15%	31%
Secundaria Incompleta	4122	23%	60%	4310	24%	55%
Secundaria Completa	4923	28%	88%	5570	31%	86%
Técnico/Tecnólogo Incompleto	253	1%	89%	198	1%	87%
Técnico/Tecnólogo Completo	862	5%	94%	692	4%	91%
Profesional Completo	779	4%	98%	299	2%	93%
Profesional Incompleta	184	1%	99%	1131	6%	99%
Posgrado	104	1%	100%	106	1%	100%
<b>Total</b>	<b>17733</b>	<b>100%</b>		<b>17733</b>	<b>100%</b>	

**Fuente:** Pruebas ICFES Saber 11º, segundo semestre de 2010

La Tabla 10 muestra por rangos el número de personas que conforman el hogar de los estudiantes, siendo los hogares con menos de cinco personas un 53%. Adicionalmente, se consideran dos variables institucionales, la primera muestra que tan sólo el 3% de los estudiantes provienen de un colegio privado y la segunda, que las jornadas a las cuales asisten más estudiantes son la mañana y la tarde, mostrando que muy pocos asisten a la jornada completa.

**Tabla 9. Proviene de colegio privado**

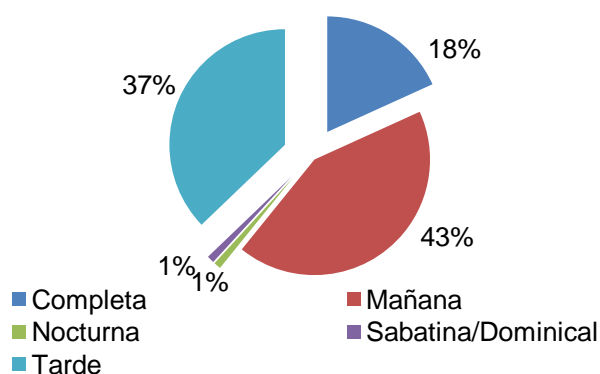
	Estudiantes	Porcentaje
<b>No</b>	17180	97%
<b>Si</b>	553	3%
<b>Total</b>	17733	100%

**Tabla 10. Personas en el hogar**

Rango	Estudiantes	Porcentaje
<b>Hasta 4 personas</b>	9473	53%
<b>De 5 a 8 personas</b>	7659	43%
<b>Más de 9 personas</b>	601	3%
<b>Total</b>	17733	100%

**Fuente:** Pruebas ICFES Saber 11°, segundo semestre de 2010

**Gráfico 6. Jornada del colegio**



**Fuente:** Pruebas ICFES Saber 11°, segundo semestre de 2010

Finalmente, en la Tabla 11 se expone la distribución de los puntajes de los estudiantes en las áreas de matemáticas y lenguaje. Con el fin de facilitar la descripción de estos datos, se presentan de acuerdo con la clasificación que hace el ICFES para las áreas del núcleo común (Bajo (0-30), Medio (30.01-70) y Alto (Más de 70.01)). Sin embargo, para la investigación realizada esta variable se tomó de manera continua.

**Tabla 11. Resultado de matemáticas y lenguaje**

Nivel	Matemáticas		Lenguaje	
	Frecuencia	Porcentaje	Frecuencia	Porcentaje
Bajo	1284	7%	156	1%
Medio	16291	92%	17562	99%
Alto	158	1%	15	0%
<b>Total</b>	<b>17733</b>	<b>100%</b>	<b>17733</b>	<b>100%</b>

**Fuente:** Pruebas ICFES Saber 11°, segundo semestre de 2010

## 6. Resultados

A continuación se presentan los principales resultados de las estimaciones realizadas. Para cada materia se realizó el ejercicio por MCO y regresión por cuantiles.

Las pruebas para el modelo de MCO se exponen en el Anexo 1, se realizaron pruebas de Varianza Inflada (VIF) y Test de Wald (significancia global). Se estimaron errores estándar robustos. El Test de varianza inflada VIF dio mayor que 10 para los primeros 4 niveles de educación de la madre, sin embargo las estimaciones de MCO dan significativas para niveles de educación por encima del cuarto, por tanto no se genera problemas con dichas variables, las demás están en el rango permitido del Test ( $VIF < 10$ ), al igual que la media de VIF.<sup>10</sup>

**Tabla 12. Estadísticas descriptivas de las variables dependientes.**

Variable	Observaciones	Media	Desviación estándar	Min	Max
Matemáticas	17733	44.43781	9.704183	0	111.55
Lenguaje	17733	46.73638	6.73268	11.11	72.22

**Fuente:** Cálculo del autor.

La Tabla 12 muestra las estadísticas descriptivas de las variables dependientes con las cuales se hace más fácil la interpretación y entendimiento de los efectos que se encuentran en las estimaciones.

### 6.1. Matemáticas MCO y cuantiles.

En la Tabla 13 se expone la estimación realizada por el método de MCO para el área de matemáticas. El método de MCO permitió, en primera instancia, conocer aquellas variables que tienen un mayor impacto en el rendimiento académico de los estudiantes. Se realizaron varios modelos teniendo en cuenta las diferentes variables explicativas, sin embargo las estimaciones no arrojaron evidencia estadística a favor de la influencia en el rendimiento de variables

---

<sup>10</sup> El Test de Factor de Incremento de Varianza (VIF) se utiliza para detectar multicolinealidad o dependencia lineal en las variables explicativas. Este se calcula como el recíproco del factor de tolerancia  $VIF = 1/(1-R^2_y)$ ;  $VIF > 10$  se consideran indicativos de multicolinealidad.

como servicio tv e internet, y otras variables, es así como el análisis se basa en el modelo que se presenta en la columna 5 el cual incluye las variables significativas del modelo en cuanto a *peer effects* y otras variables explicativas como edad, género, trabaja, educación de los padres, ingresos entre otras.

Al revisar los efectos de las variables de *peer effects* se pueden distinguir dos clases de efectos teniendo en cuenta los datos con los que se cuenta. El *efecto endógeno*, en este caso se ve reflejado en las variables que miden el rendimiento promedio de los compañeros de curso, es decir promedio matemáticas y promedio lenguaje. El *efecto de contexto*, que se entiende como influencia que tiene sobre el desempeño de un estudiante las características de sus compañeros de clase que para el caso de este trabajo se mide a través de los niveles promedio de educación de los padres y nivel de ingresos.

La estimación por MCO muestra una clara evidencia estadística a favor de **efectos de pares endógenos**, es decir, que el rendimiento promedio de los compañeros de grupo afecta el desempeño del estudiante. Se observa como el desempeño promedio en matemáticas afecta en mayor medida que el desempeño promedio en lenguaje, pero la diferencia entre estas sólo es de 0.057 puntos. Esta diferencia es razonable porque se trata de un área que requiere un mayor número de conocimientos específicos, sin embargo, el ser tan mínima evidencia la importancia que tiene en el aprendizaje de cualquier área del conocimiento las competencias de comprensión lectora y análisis de textos. El efecto de pares positivo se fundamenta en que el estar en clase con compañeros de alto nivel genera un efecto de transferencia de conocimientos, a su vez que se generan otro tipo de externalidades positivas como mayor concentración en clase, buena disposición para estudiar, y apoyo y motivación de los compañeros. Al revisar la magnitud del efecto podemos ver que un aumento de un punto en el promedio en matemáticas aumenta el promedio del estudiante en 0.389 es decir, aumenta 4% de desviaciones estándar, este efecto no es para nada despreciable teniendo en cuenta que existen estudiantes con desempeños muy por debajo del promedio donde el efecto sería mucho mayor, así mismo que la curva de aprendizaje cuando se habla de educación es bastante elevada. Por su parte, el efecto de par que genera el



promedio de lenguaje es de 3% desviaciones estándar un poco menor que el de matemáticas, lo cual como se explicó anteriormente es razonable por la especificidad que requiere el área de matemáticas.

Sin embargo la estimación no arroja evidencia de la influencia en el desempeño de **efectos de pares de contexto**, medido por variables como nivel de ingreso y estudios promedios de los padres, lo anterior puede estar sustentado en la homogeneidad que presentan los datos de la base. Al revisar detenidamente el Gráfico 5 y la Tabla 8 se observa que aproximadamente 90% de los estudiantes se concentran en ingresos inferiores a 3 SMLV y los padres se concentran en la misma proporción en tener menos de educación secundaria. Esta homogeneidad puede suavizar el efecto que tengan estas variables promedio en el desempeño, ya que los estudiantes tendrán compañeros que características muy similares de este tipo.

Por su parte, al revisar las otras variables incluidas en el estudio se tiene que variables personales como género hombre (+), trabajo (-) y edad (-) resultaron ser estadísticamente significativas. Al revisar las variables socioeconómicas se encuentra que por cada persona adicional en el hogar se disminuye los resultados en 0.075 puntos. En cuanto a recursos y servicios que tiene el estudiante, se destaca que el tener computador impacta positivamente los resultados, pero, que tener DVD y disminuye el desempeño, esto puede ser explicado porque el computador es una herramienta que facilita el proceso de aprendizaje, mientras que el DVD puede ser visto como elemento distractor de este proceso.

Para la variable ingresos mensuales se observa que los estudiantes con ingresos familiares mayores comparados con el nivel de mínimo ingreso se ven afectados positivamente en el rendimiento, y a medida que aumenta el ingreso, también lo hace la magnitud de este impacto, lo cual ha sido demostrado por diversos estudios. Otra variable importante es el nivel de educación de los padres, para ésta se puede concluir que sólo a partir de que los padres culminan su educación secundaria esta variable comienza a tener significancia en el desempeño de sus hijos, lo cual es bastante coherente ya que se está

evaluando el desempeño en el último grado de la secundaria. En cuanto a variables de tipo institucional muestra que el hecho de que un estudiante provenga de un colegio privado afecta positivamente su desempeño en la prueba, y el pertenecer a instituciones con jornada nocturna o sabatina afecta negativamente el desempeño en relación con la jornada completa.

**Tabla 13. Estimación por MCO para el área de matemáticas.**

Variables	Variable dependiente: Resultado matemáticas				
	1	2	3	4	5
<b>Género</b>	3.031***	3.031***	3.023***	3.020***	3.022***
	(0.141)	(0.141)	(0.140)	(0.140)	(0.141)
<b>Edad</b>	-0.789***	-0.791***	-0.791***	-0.790***	-0.787***
	(0.0595)	(0.0595)	(0.0595)	(0.0594)	(0.0634)
<b>Trabaja</b>	-0.714*	-0.711*	-0.709*	-0.706*	-0.728**
	(0.369)	(0.369)	(0.369)	(0.369)	(0.361)
<b>Tiene computador</b>	0.880***	0.856***	0.851***	0.849***	0.845***
	(0.201)	(0.155)	(0.155)	(0.154)	(0.153)
<b>Reproductor DVD</b>	-0.666***	-0.706***	-0.707***	-0.707***	-0.702***
	(0.151)	(0.149)	(0.149)	(0.149)	(0.149)
<b>Servicio TV</b>	-0.258				
	(0.164)				
<b>Internet</b>	0.0161				
	(0.199)				
<b>Ingresos mensuales 2</b>	0.471***	0.445**	0.449**	0.446**	0.404**
	(0.179)	(0.179)	(0.179)	(0.179)	(0.170)
<b>Ingresos mensuales 3</b>	0.474*	0.431*	0.438*	0.430*	0.421*
	(0.251)	(0.249)	(0.249)	(0.248)	(0.243)
<b>Ingresos mensuales 4</b>	1.538***	1.494***	1.502***	1.507***	1.516***
	(0.394)	(0.392)	(0.391)	(0.391)	(0.403)
<b>Ingresos mensuales 5</b>	1.823**	1.777*	1.769*	1.781*	1.811*
	(0.915)	(0.914)	(0.913)	(0.913)	(1.033)
<b>Ingresos mensuales 6</b>	5.240***	5.197***	5.143**	5.139**	5.189**
	(2.018)	(2.017)	(2.017)	(2.017)	(2.143)
<b>Ingresos mensuales 7</b>	-0.0603	-0.101	-0.141	-0.122	-0.0415
	(2.354)	(2.354)	(2.353)	(2.353)	(3.347)
<b>Nivel padre 1</b>	0.115	0.105	0.0864	0.0792	0.0612
	(0.401)	(0.401)	(0.401)	(0.399)	(0.388)
<b>Nivel padre 2</b>	0.0312	0.0242	0.0135	0.00180	-0.0114
	(0.401)	(0.401)	(0.401)	(0.398)	(0.390)
<b>Nivel padre 3</b>	0.621	0.607	0.580	0.559	0.534
	(0.396)	(0.396)	(0.394)	(0.392)	(0.386)
<b>Nivel padre 4</b>	0.480	0.465	0.424	0.411	0.395
	(0.396)	(0.396)	(0.394)	(0.392)	(0.387)
<b>Nivel padre 5</b>	2.806***	2.794***	2.748***	2.737***	2.717***
	(0.687)	(0.687)	(0.686)	(0.685)	(0.714)
<b>Nivel padre 6</b>	2.641***	2.622***	2.556***	2.552***	2.539***
	(0.496)	(0.496)	(0.495)	(0.493)	(0.497)

Nivel padre 7	3.647***	3.639***	3.569***	3.566***	3.552***
	(0.773)	(0.773)	(0.772)	(0.771)	(0.869)
Nivel padre 8	2.451***	2.430***	2.345***	2.342***	2.325***
	(0.512)	(0.512)	(0.510)	(0.508)	(0.526)
Nivel padre 9	4.801***	4.787***	4.697***	4.716***	4.703***
	(1.004)	(1.004)	(1.002)	(1.001)	(1.180)
Nivel madre 1	-0.00491	-0.0176	-0.0168	-0.0242	-0.0449
	(0.628)	(0.628)	(0.628)	(0.610)	(0.575)
Nivel madre 2	0.793	0.776	0.777	0.760	0.736
	(0.626)	(0.626)	(0.626)	(0.609)	(0.577)
Nivel madre 3	0.590	0.570	0.569	0.549	0.519
	(0.621)	(0.621)	(0.621)	(0.603)	(0.569)
Nivel madre 4	1.049*	1.031*	1.031*	1.012*	0.990*
	(0.621)	(0.621)	(0.621)	(0.603)	(0.571)
Nivel madre 5	2.869***	2.852***	2.862***	2.854***	2.838***
	(0.807)	(0.807)	(0.807)	(0.793)	(0.824)
Nivel madre 6	2.339***	2.321***	2.308***	2.295***	2.277***
	(0.675)	(0.674)	(0.674)	(0.657)	(0.632)
Nivel madre 7	3.627***	3.615***	3.563***	3.546***	3.515***
	(0.893)	(0.893)	(0.893)	(0.880)	(0.830)
Nivel madre 8	1.641**	1.623**	1.621**	1.613**	1.603**
	(0.713)	(0.713)	(0.713)	(0.697)	(0.685)
Nivel madre 9	3.061***	3.038***	3.040***	3.040***	2.976**
	(1.098)	(1.098)	(1.098)	(1.087)	(1.176)
Proviene privado	1.501***	1.499***	1.468***	1.486***	1.516***
	(0.396)	(0.396)	(0.396)	(0.395)	(0.447)
Personas hogar	-0.0751*	-0.0762*	-0.0760*	-0.0756*	-0.0756*
	(0.0406)	(0.0406)	(0.0406)	(0.0406)	(0.0400)
Jornada 2	0.346*	0.339*	0.367*	0.357*	0.324*
	(0.197)	(0.197)	(0.195)	(0.194)	(0.193)
Jornada 3	-1.423**	-1.424**	-1.366*	-1.398**	-1.335*
	(0.713)	(0.713)	(0.712)	(0.711)	(0.727)
Jornada 4	-1.443**	-1.465**	-1.424**	-1.434**	-1.492**
	(0.699)	(0.699)	(0.698)	(0.696)	(0.650)
Jornada 5	0.175	0.169	0.182	0.175	0.133
	(0.198)	(0.198)	(0.197)	(0.197)	(0.195)
Promedio ingresos 2	-0.363	-0.371	-0.388	-0.471	
	(0.338)	(0.338)	(0.336)	(0.318)	
Promedio Ingresos 3	0.0829	0.0640	-0.0704	-0.0501	
	(0.441)	(0.441)	(0.436)	(0.411)	
Promedio ingresos 4	-9.165	-9.202	-4.766	-4.529	
	(9.184)	(9.185)	(6.431)	(6.420)	
Promedio nivel padre 1	-0.836	-0.710	1.100		
	(6.462)	(6.461)	(2.574)		
Promedio nivel padre 2	-0.229	-0.108	0.489		
	(6.460)	(6.459)	(2.247)		
Promedio nivel padre 3	-0.329	-0.226	0.320		
	(6.461)	(6.461)	(2.253)		
Promedio nivel padre 4	-0.549	-0.450	0.388		
	(6.463)	(6.462)	(2.259)		
Promedio nivel padre 5	-1.068	-0.966	0.703		
	(6.474)	(6.473)	(2.288)		

<b>Promedio nivel padre 6</b>	-2.479	-2.360			
	(6.634)	(6.634)			
<b>Promedio nivel padre 8</b>	9.847	9.961			
	(14.43)	(14.43)			
<b>Promedio nivel madre 1</b>	1.845	1.727			
	(6.824)	(6.824)			
<b>Promedio nivel madre 2</b>	0.751	0.567			
	(6.842)	(6.841)			
<b>Promedio nivel madre 3</b>	0.652	0.456			
	(6.847)	(6.846)			
<b>Promedio nivel madre 4</b>	0.845	0.647			
	(6.851)	(6.850)			
<b>Promedio nivel madre 5</b>	1.413	1.217			
	(6.863)	(6.862)			
<b>Promedio matemáticas</b>	0.382***	0.382***	0.380***	0.380***	0.389***
	(0.0329)	(0.0329)	(0.0327)	(0.0319)	(0.0325)
<b>Promedio lenguaje</b>	0.337***	0.340***	0.330***	0.338***	0.332***
	(0.0485)	(0.0484)	(0.0482)	(0.0471)	(0.0462)
<b>Constante</b>	22.01***	21.97***	22.47***	22.54***	22.08***
	(2.941)	(2.941)	(2.924)	(2.017)	(2.032)
<b>Observaciones</b>	17733	17733	17733	17733	17733
<b>R-cuadrado</b>	0.133	0.133	0.133	0.132	0.132

Errores estándar en  
paréntesis

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Fuente:** Cálculo del autor.

Con respecto a los *peer effects* la regresión por cuantiles para el área de matemáticas expuesta en el Anexo 2, evidencia que tener compañeros con mejores rendimientos no afecta a los estudiantes que se ubican en el cuantil más bajo, y sólo comienza a afectar a partir del cuantil 25. Se destaca que el rendimiento de pares en lenguaje comienza a tener efectos positivos y significativos en el rendimiento de matemáticas a partir del cuantil 50 pero con rendimientos decrecientes.

La regresión por cuantiles permite concluir con respecto a la variable género, que el ser mujer u hombre no afecta en el cuantil más bajo, es decir que para los estudiantes con menores resultados no es significativa esta variable, lo mismo ocurre para la variable edad. En cuanto a los recursos del estudiante se observa que el tener computador afecta positivamente los cuantiles más altos, lo cual puede dar evidencia de que en los cuantiles más bajos o no se tiene la herramienta o no se utiliza ésta de manera adecuada. Con referencia a los

ingresos, se destaca en la regresión por cuantiles, que para estudiantes del ingreso 5 y 6 la estimación por mínimos cuadrados subestima los efectos que tiene esta variable sobre los cuantiles de rendimiento más bajo. Por ejemplo, para los cuantiles q5 y q25 tener un ingreso mensual familiar de nivel 6 aumenta en alrededor 10 a 15 puntos el desempeño con respecto a los que perciben un ingreso menor a un salario mínimo, mientras que por MCO este aumento se valora en 5.1 puntos.

## 6.2. Lenguaje MCO y cuantiles.

En la Tabla 14 se muestra que los efectos de pares en el área de lenguaje tienen resultados similares que los encontrados para el área de matemáticas. Los **efectos de pares de contexto** resultan no ser significativos (nivel educación e ingreso promedios) lo cual es razonable partiendo de que se tienen los mismos datos que para la estimación por matemáticas. De nuevo el **efecto de pares endógeno** resulta ser significativo, el tener compañeros de clase con habilidades para el lenguaje afecta positivamente en 0.41 puntos, es decir 6% de desviaciones estándar, mientras que contar con compañeros con buenos rendimientos en matemáticas afecta positivamente en tan sólo 0.11 puntos el rendimiento en lenguaje, es decir 2% de desviaciones estándar. Comparando las estimaciones para ambas áreas sobresale que el efecto de pares de matemáticas y lenguaje, influye en la misma magnitud para obtener buenos rendimientos en el área de matemáticas, contrario a lo que ocurre cuando la variable a explicar es lenguaje. Así, se muestra evidencia a favor de que el desempeño en lenguaje se ve afectado en mayor medida por contar con compañeros con buenos niveles promedio en lenguaje.

Otros resultados expuestos en la Tabla 14 muestran evidencia que la dirección del impacto de las diferentes variables explicativas sobre el rendimiento académico es similar a la estimación en matemáticas para la mayoría de variables, no obstante, se muestra que la magnitud de la influencia de las variables es mayor en el área de matemáticas que en la de lenguaje, excepto para variable como la jornada de la noche. Se destaca que variables como servicio de televisión e internet aparecen como variables significativas, sin

embargo la televisión aparece relacionada negativamente mientras que el servicio de internet aparece positivamente, esto último puede verse explicado porque el acceso a esta herramienta podría estar relacionado con la capacidad de adquirir un vocabulario más extenso, interactuar con otras personas, entre otros.

**Tabla 14. Estimación por MCO para el área de lenguaje.**

Variables	Variable dependiente: Resultado lenguaje			
	1	2	3	4
<b>Género</b>	0.326*** (0.0995)	0.328*** (0.0994)	0.327*** (0.0992)	0.327*** (0.0994)
<b>Edad</b>	-0.354*** (0.0421)	-0.353*** (0.0421)	-0.351*** (0.0421)	-0.354*** (0.0448)
<b>Trabaja</b>	-0.406 (0.261)	-0.400 (0.261)	-0.401 (0.261)	-0.401 (0.255)
<b>Tiene computador</b>	0.354** (0.142)	0.352** (0.142)	0.345** (0.142)	0.348** (0.145)
<b>Reproductor DVD</b>	-0.318*** (0.107)	-0.317*** (0.107)	-0.316*** (0.107)	-0.317*** (0.107)
<b>Servicio TV</b>	-0.267** (0.116)	-0.271** (0.116)	-0.291** (0.115)	-0.291** (0.116)
<b>Internet</b>	0.280** (0.141)	0.278** (0.141)	0.278** (0.141)	0.274* (0.143)
<b>Ingresos mensuales 2</b>	0.307** (0.127)	0.306** (0.127)	0.306** (0.127)	0.316** (0.123)
<b>Ingresos mensuales 3</b>	0.628*** (0.178)	0.627*** (0.178)	0.622*** (0.177)	0.613*** (0.175)
<b>Ingresos mensuales 4</b>	1.312*** (0.279)	1.308*** (0.279)	1.316*** (0.279)	1.303*** (0.282)
<b>Ingresos mensuales 5</b>	0.716 (0.647)	0.715 (0.647)	0.724 (0.647)	0.701 (0.732)
<b>Ingresos mensuales 6</b>	2.953** (1.428)	2.934** (1.428)	2.940** (1.428)	2.910* (1.556)
<b>Ingresos mensuales 7</b>	4.143** (1.667)	4.135** (1.666)	4.145** (1.666)	4.099** (2.049)
<b>Nivel padre 1</b>	-0.189 (0.284)	-0.190 (0.284)	-0.201 (0.282)	-0.196 (0.281)
<b>Nivel padre 2</b>	-0.120 (0.284)	-0.125 (0.284)	-0.145 (0.282)	-0.141 (0.281)
<b>Nivel padre 3</b>	0.0979 (0.280)	0.0838 (0.279)	0.0425 (0.278)	0.0495 (0.277)
<b>Nivel padre 4</b>	0.213 (0.280)	0.197 (0.279)	0.167 (0.278)	0.169 (0.276)
<b>Nivel padre 5</b>	1.650*** (0.486)	1.635*** (0.486)	1.608*** (0.485)	1.613*** (0.497)
<b>Nivel padre 6</b>	1.364*** (0.351)	1.343*** (0.350)	1.320*** (0.349)	1.323*** (0.348)

Nivel padre 7	1.975***	1.956***	1.936***	1.938***
	(0.547)	(0.547)	(0.546)	(0.614)
Nivel padre 8	1.356***	1.335***	1.311***	1.307***
	(0.363)	(0.361)	(0.360)	(0.365)
Nivel padre 9	1.803**	1.775**	1.770**	1.773**
	(0.711)	(0.710)	(0.709)	(0.706)
Nivel madre 1	0.242	0.253	0.213	0.218
	(0.444)	(0.444)	(0.432)	(0.428)
Nivel madre 2	0.159	0.163	0.102	0.107
	(0.444)	(0.443)	(0.431)	(0.427)
Nivel madre 3	0.412	0.417	0.343	0.348
	(0.440)	(0.439)	(0.427)	(0.422)
Nivel madre 4	0.695	0.699	0.625	0.630
	(0.440)	(0.440)	(0.427)	(0.422)
Nivel madre 5	1.524***	1.532***	1.472***	1.474***
	(0.571)	(0.571)	(0.561)	(0.558)
Nivel madre 6	1.573***	1.575***	1.508***	1.512***
	(0.478)	(0.477)	(0.466)	(0.465)
Nivel madre 7	2.311***	2.311***	2.243***	2.252***
	(0.632)	(0.632)	(0.623)	(0.603)
Nivel madre 8	1.221**	1.232**	1.170**	1.172**
	(0.505)	(0.504)	(0.493)	(0.499)
Nivel madre 9	1.755**	1.763**	1.711**	1.736**
	(0.777)	(0.777)	(0.770)	(0.847)
Proviene privado	1.203***	1.201***	1.223***	1.210***
	(0.281)	(0.280)	(0.280)	(0.269)
Personas hogar	-0.0561*	-0.0559*	-0.0550*	-0.0548*
	(0.0287)	(0.0287)	(0.0287)	(0.0286)
Jornada 2	0.195	0.197	0.189	0.210
	(0.139)	(0.138)	(0.138)	(0.137)
Jornada 3	-2.049***	-2.046***	-2.088***	-2.116***
	(0.505)	(0.504)	(0.503)	(0.525)
Jornada 4	-1.447***	-1.428***	-1.433***	-1.409***
	(0.495)	(0.494)	(0.493)	(0.456)
Jornada 5	-0.0831	-0.0851	-0.0999	-0.0775
	(0.140)	(0.139)	(0.139)	(0.138)
Promedio ingresos 2	0.321	0.303	0.127	
	(0.239)	(0.238)	(0.225)	
Promedio Ingresos 3	0.0546	0.0196	-0.0757	
	(0.312)	(0.309)	(0.291)	
Promedio ingresos 4	-5.882	-3.368	-3.372	
	(6.502)	(4.552)	(4.545)	
Promedio nivel padre 1	-4.129	0.857		
	(4.574)	(1.822)		
Promedio nivel padre 2	-4.279	0.239		
	(4.573)	(1.591)		
Promedio nivel padre 3	-4.391	-0.187		
	(4.574)	(1.595)		
Promedio nivel padre 4	-4.481	-0.0764		
	(4.575)	(1.600)		
Promedio nivel padre 5	-4.484	0.0961		
	(4.583)	(1.620)		

<b>Promedio nivel padre 6</b>	-4.666			
	(4.697)			
<b>Promedio nivel padre 8</b>	0.829			
	(10.21)			
<b>Promedio nivel madre 1</b>	4.888			
	(4.831)			
<b>Promedio nivel madre 2</b>	4.490			
	(4.844)			
<b>Promedio nivel madre 3</b>	4.136			
	(4.847)			
<b>Promedio nivel madre 4</b>	4.306			
	(4.850)			
<b>Promedio nivel madre 5</b>	4.494			
	(4.859)			
<b>Promedio matemáticas</b>	0.121***	0.119***	0.114***	0.111***
	(0.0233)	(0.0231)	(0.0226)	(0.0223)
<b>Promedio lenguaje</b>	0.408***	0.406***	0.415***	0.415***
	(0.0343)	(0.0341)	(0.0333)	(0.0332)
<b>Constante</b>	27.17***	27.32***	27.31***	27.55***
	(2.082)	(2.070)	(1.429)	(1.438)
<b>Observaciones</b>	17733	17733	17733	17733
<b>R-cuadrado</b>	0.097	0.097	0.097	0.097

Errores estándar en  
paréntesis

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Fuente:** Cálculo del autor.

La estimación por cuantiles para el área de lenguaje se expone en el Anexo 3. En él se destaca que la estimación por MCO subestima el valor del efecto que tiene el que un estudiante llegue de un colegio privado sobre el cuantil más bajo, es decir la estimación por cuantiles revela que si un estudiante que se ubicase en el cuantil más bajo proviniese del sector privado, esto aumentaría en 4.18 puntos su rendimiento en la prueba de lenguaje. A medida que nos alejamos a cuantiles más altos el efecto de esta sobre el rendimiento disminuye.

En cuanto a los *peer effects*, se muestra que la estimación por MCO es positiva y significativa, además las estimaciones por cuantiles son también positivas y significativas y relativamente constantes con valores cercanos al efecto medio. Dado lo anterior, se puede concluir que en este caso el efecto sobre el promedio provee un resumen representativo del efecto del rendimiento de los pares sobre toda la distribución.



### 6.3. Modelos lineales jerárquicos (HLM).

Con el fin de tener en cuenta la estructura anidada de los datos que se presenta al estudiar problemas relacionados con la educación, y tratar de corregir una de las limitaciones que plantea la estimación por MCO, los cuales asumen independencia entre los errores estándar, este aparte muestra la simulación de un modelo jerárquico lineal sencillo, tomando como base algunos de los resultados obtenidos mediante la estimación de MCO para determinar la posibilidad de incluir este tipo de análisis en el estudio de *peer effects*.

Se estiman los modelos expuestos en la sección 4.2, es decir la primera columna corresponde al modelo nulo o incondicional, la segunda columna corresponde al modelo con una variable explicativa a primer nivel (género), en este caso se tomo está variable porque era la que a primer nivel generaba influencias en mayor magnitud en el rendimiento académico.

**Tabla 15. Modelo Jerárquico lineal para matemáticas y lenguaje**

	Matemáticas		Lenguaje	
Modelo	Modelo 1 nulo	Modelo 2	Modelo 1 nulo	Modelo 2
<b>Efectos fijos</b>				
$\gamma_{00}$	44.25482*** (0.13856)	42.72983*** (0.15228)	46.59058*** (0.09371)	46.36337*** (0.10626)
$\gamma_{01}$				
$\gamma_{10}$		3.388038*** (0.15228)		0.5043404*** (0.10706)
<b>Varianzas y covarianzas</b>				
$\tau_{00}$	8.252746*** (0.65361)	9.088795*** (0.11568)	3.694915*** (0.29776)	3.819457***
$\tau_{11}$				
$\tau_{01}$				
$\sigma^2$	85.74087*** (0.92576)	83.16836*** (0.04924)	41.52285 *** (0.44826)	41.43901*** (0.44742)
$\rho$ (CCI)	9.62%	9.85 %	8.17%	8.43%

Criterios de información del modelo				
<b>AIC</b>	<b>130052.2</b>	<b>129570.3</b>	<b>117161.6</b>	<b>117144.2</b>
<b>BIC</b>	<b>130075.5</b>	<b>129601.4</b>	<b>117185</b>	<b>117175.4</b>

**Fuente:** Cálculo del autor.

*Para comprobar la necesidad de utilizar los modelos HLM se calcula el ICC ( $\rho$ ) el cual se explicó en la sección 4.2, por tanto retomando:*

*Si  $\rho = 0$ , colegios o grupos homogéneos.*

*Si  $\rho = 1$ , no hay variabilidad al interior de los grupos, estudiantes iguales.*

Si se encuentra un  $\rho > 0$ , se puede concluir que es viable la estimación por HLM con el fin de capturar la variabilidad entre grupos. En la Tabla 15 se observa que el ICC para matemáticas es en el modelo nulo de 9.62%, que para este caso se refiere a que la variabilidad al interior de los grupos de los colegios es de 9.62%, mientras que entre los grupos de los colegios es de 90.38%. Al incluir una variable independiente a nivel 1, encontramos que la variabilidad al interior de los grupos aumenta un poco, lo cual indica que sí se deben tener en cuenta variables con las cuales se pueda explicar esta variabilidad al interior de los grupos, es decir efectos de pares. En lenguaje la variabilidad es un poco menor, pero el análisis es similar. Por su parte, los criterios de AIC y BIC se van haciendo más pequeños cada vez que se introducen variables explicativas, en este sentido también revela la necesidad de incluir más covariables tanto a nivel de estudiantes como a nivel de grupo con el fin de capturar mejor los efectos. Este estudio deja abierta la posibilidad de mejorar la estimación de MCO por medio de modelos HLM donde se tenga en cuenta la estructura jerárquica del proceso, adicionalmente se puede incluir un tercer nivel que corresponda a variables de contexto como violencia de la zona, entre otras.

## 7. Conclusiones

La literatura internacional destaca dentro de los factores determinantes de la calidad o desempeño de los estudiantes el efecto que ejercen los compañeros de clase sobre el rendimiento de cada estudiante (*peer effects*). En Colombia esto no ha sido analizado aun, dada la falta de datos de rendimiento escolar a nivel de grupo. En esta investigación se realizó un estudio sobre la existencia de efectos de pares en el rendimiento académico de los estudiantes de grado 11 de los colegios oficiales de la ciudad de Medellín con base en los resultados de las pruebas Saber para el segundo semestre de 2010 para las áreas de lenguaje y matemáticas. Este trabajo presenta una primera evidencia a favor de la existencia de efectos de pares para los estudiantes de IE oficiales de la ciudad de Medellín.

Adicional a las variables que determinaban efectos de pares se tuvieron en cuenta variables de tipo personal, socioeconómico e institucional con el fin de no desconocer la importancia que tienen estas dimensiones en los logros alcanzados. El análisis se realizó a partir de la estimación de tres modelos econométricos: un primer modelo a través de la metodología de Mínimos Cuadrados Ordinarios (MCO) el cual permitió conocer cuáles eran las variables que más explican el rendimiento académico; un segundo modelo, por medio de la estimación de regresión por cuantiles con el que se pretendía conocer cuáles de las variables explicativas tienen efectos diferentes a través de los cuantiles de la distribución de los resultados en las pruebas y por último una aproximación a los modelos jerárquicos lineales.

Para la determinación de *peer effects* se estimaron varios modelos en los cuales se incluía diferentes variables que podría determinar efectos de pares al interior de las aulas, tales como: nivel de educación promedio de los padres, nivel de ingresos promedios de la familia, promedio de resultado en el área de matemáticas y promedio de resultado en el área de lenguaje. Los resultados arrojan evidencia estadística de que en los colegios oficiales de la ciudad de Medellín existen efectos de *pares endógenos*, es decir aquel que depende del desempeño de los compañeros, tanto para lenguaje como para matemáticas.

Sin embargo se presenta ausencia de efectos de *pares de contexto*. Una explicación a esta ausencia se puede encontrar al analizar la composición de la base de datos estudiada, la mayor parte de la población se concentra en los niveles de ingresos inferior a 3 SMLV y la educación de los padres se concentra en alrededor de un 90% en secundaria completa, con lo cual se puede inferir que el cálculo de los promedios de las variables socioeconómicas consideradas en las aulas, serían uniformes entre estudiantes de un mismo grupo. Por su parte la regresión por cuantiles, evidencia la importancia que tiene en los diferentes niveles de la distribución del resultado los efectos de pares para las dos áreas.

Los resultados muestran evidencia estadística para aceptar la hipótesis planteada en este trabajo: existen efectos de pares en el rendimiento académico de los estudiantes de grado 11° en las instituciones educativas oficiales de la ciudad de Medellín. A partir de esto se generan consideraciones en cuanto a la importancia que tiene para el desempeño de un área específica el contar con compañeros que cuenten con un desarrollo integral. Por ejemplo, se destaca que para obtener resultados positivos en el área de matemáticas influye casi en la misma magnitud el contar con compañeros de clase con buenos resultados tanto para matemáticas como para lenguaje. Cabe resaltar que en el resultado en el área de lenguaje no influye de la misma manera el resultado promedio de matemáticas.

Este efecto de pares positivo se fundamenta en que el estar en clase con compañeros de alto nivel genera un efecto de transferencia de conocimientos, a su vez que se forma otro tipo de externalidades positivas como mayor concentración en clase, buena disposición para estudiar, y apoyo y motivación de los compañeros. El anterior resultado permite generar reflexiones referentes a la gestión escolar, este sugiere que las instituciones deben implementar mecanismos encaminados a potencializar la heterogeneidad al interior de las aulas por medio de estrategias que permiten reunir en los salones estudiantes tanto de desempeños altos como bajos, de manera que se maximicen las externalidades derivadas de los efectos de pares existentes. Así como apoyar económicamente o crear incentivos para estudiantes brillantes de manera que

su presencia en las aulas motive otros compañeros. Estas estrategias tienen la ventaja de que no requieren grandes inversiones como lo hacen otro tipo de políticas, y además pueden ser puestas en marchas por cada rector o institución lo cual minimiza en gran medida los trámites que se pueden generar. Si bien es cierto que se puede llegar a pensar que este tipo de políticas pueden generar el efecto contrario, es decir bajar el rendimiento de los alumnos mejores, se encuentran estudios como los de Kang (2007) y Sund (2009) que respaldan que en definitiva escoger una política de mezcla (mezclar estudiantes de promedios altos con promedios bajos) es mucho mejor y trae mayores beneficios en términos de rendimiento promedio, que el tener políticas de agrupación (agrupar los alumnos de acuerdo a su rendimiento), ya que el resultado de esta última política por lo general ocasiona mayor daño al disminuir los resultados promedios de una institución.

La estimación de modelos HLM permitió concluir que existe un Coeficiente de Correlación Intraclase (CCI) significativo el cual muestra evidencia a favor de la necesidad de realizar estudios de *peer effects* teniendo en cuenta la estructura jerárquica del proceso escolar con el fin de conocer hasta qué punto la estimación por MCO puede estar sobreestimando los resultados al no tener en cuenta la estructura anidada de los datos. Lo anterior, deja entre abierto un gran camino por recorrer en el estudio de las variables de *peer effect* en nuestro país. Las próximas investigaciones que se generen en torno a este tema podrían ampliar los resultados teniendo en cuenta variables que permitan medir los *efectos de pares correlacionados*, por medio de la inclusión de variable que midan ambiente similares entre estudiantes como lo son variables de entorno social (índices de violencia) o clima escolar, entre otros. También podrían incluirse estudios que se amplíen a colegios tanto oficiales como privados explorando la heterogeneidad que pueda darse en este tipo de colegios y en los cuales los efectos de pares de contexto como el nivel promedio de educación de los padres y de ingresos pueden resultar significativos. Todo lo anterior se podría realizar aprovechando la metodología de HLM ampliando el estudio a tres niveles, incluyendo en un primer nivel las variables de cada estudiante, en un segundo nivel las variables de *peer effects* y un último en el cual se mida el entorno de los estudiantes o la institución

## Bibliografía

- Alexander, L. y Simmons, J. (1975). "The Determinants of School Achievement in Developing Countries: the educational production function". En: Cuadernos de Trabajo. No. 201. Washington: Banco Mundial.
- Becker, Gary S (1964). "Human Capital". New York: Columbis Univ. Press (For Nat Bur. Econ. Res)
- Bonesrønning, Hans (2006). "Peer group effects in education production: Is it about congestion?". Department of Economics, Norwegian University of Science and Technology. The Journal of Socio-Economics pp. 328–342
- Cepeda, E. (2005). "Factores asociados al logro cognitivo en matemáticas". Revista de Educación, Número. 336 (2005), pp. 503-514. Recuperado de: [http://www.revistaeducacion.mec.es/re336/re336\\_26.pdf](http://www.revistaeducacion.mec.es/re336/re336_26.pdf)
- Coleman, J. S. et al (1966). "Equality of Educational Opportunity. Washington: US Government Printing Office".
- Correa, John Jairo (2004). "Determinantes del Rendimiento Educativo de los Estudiantes de Secundaria en Cali: un análisis multinivel". En: Revista Sociedad y Economía. No. 6. pp. 81-105.
- Gaviria, Alejandro y Barrientos, Jorge. (2001a). "Calidad de la Educación y Rendimiento Académico en Bogotá". En: Coyuntura Social, Fedesarrollo. No. 24. pp. 111-126.
- Gaviria, Alejandro y Barrientos, Jorge H. (2001b). "Determinantes de la calidad de la educación en Colombia". En: Archivos de Economía, Departamento Nacional de Planeación. No. 159.

- Gaviria, Alejandro y Barrientos, Jorge H. (2008). "Calidad de la educación pública y logro académico en Medellín 2004-2006. Una aproximación por regresión intercuartil". En: Lecturas de Economía, Universidad de Antioquia.No. 68. pp 121-144.
- Hanushek EA. (1979). "Conceptual and empirical issues in the estimation of educational production functions". Journal of Human Resources14. Pp. 351–388.
- Hanushek, Eric (1989). "The impact of Differential Expenditures on School Performance" ,Educational Researcher, Vol.18, N°4.
- Hanushek, E. A., Kain, J. F., Markman, L. M., & Rivkin, S. G. (2003). "Does peer ability affect achievement" Journal of Applied Econometrics, 18, pp. 527–544.
- Hattie, J. (2002). "Classroom composition and peer effects". School of Education, University of Auckland. International Journal of Educational Research 37 pp. 449–481
- Hedges, Larry, LAINE, Richard y Rob GREENWALD (1994). "Does Money Matter? A Meta- Analysis of Studies of Effects of Differential School Inputs on Student Outcomes". Educational Researcher, Vol.23, N°3.
- Henderson, J., P. Miezkowski y Y. Sauvageau, (1978), "Peer Group Effects and Educational Production Functions". Journal of Public Economics, vol. 10.
- Instituto Colombiano para la Evaluación de la Educación – ICFES-. Bases de datos pruebas Saber [ftp://ftp.icfes.gov.co/](http://ftp.icfes.gov.co/) (16 de junio de de 2011).

- Instituto Colombiano para la Evaluación de la Educación – ICFES-. Guía de orientación Examen de Estado (2010). <http://www.icfes.gov.co/> (17 de junio de 2011)
- Instituto Colombiano para la Evaluación de la Educación – ICFES-. <http://www.icfesinteractivo.gov.co/> (14 de junio de 2011)
- Jencks, C. S. (1972). “Inequality: a Reassessment of the Effect of Family and Schooling in America”. En: New York: Basic Books
- Kang, Changhui (2007). “Classroom peer effects and academic achievement: Quasi-randomization evidence from South Korea” . Journal of Urban Economics 61 (2007) 458–495
- Koenker, R. y D’Orey, V., 1987, “Computing regression quantiles, Applied Statistics”, 36, 383-393.
- Lazear EP. (2001). “Educational production”. Quarterly Journal of Economics 116Pp. 777–803.
- Lefgren,L. (2004). “Educational peer effects and the Chicago public schools”. Department of Economics, Brigham Young University. Journal of Urban Economics 56 pp. 169–191
- McEwan, Patrick J (2003). “Peer effects on student achievement: evidence from Chile”. Wellesley College, Department of Economics. Economics of Education Review 22 pp. 131–141
- Mankiw, N. G., Romer, D., Weil, D. N. (1992), “A Contribution to the Empirics of Economic Growth”, Quaterly Journal of Eonomis, Vol. 107, may, pp. 407-437.
- Manski, C. (1995) “Identifications Problems in the Social Sciences,”Cambridge, MA: Harvard University Press.



- Martinez, A (2011). Modelos Lineales Jerárquicos. Taller introductorio: seminario de investigación en calidad educativa ICFES 2011. Notas de taller. Noviembre de 2011.
- Robertson D. & Symons J (2003). "Do Peer Groups Matter? Peer Group versus Schooling Effects on Academy Attainment". *Economica* 70 pp. 31-53
- Schindler, B. (2003). "Educational Peer Effects Quantile Regression Evidence from Denmark with PISA 2000 data" AKF. Institute of Local Government Studies Copenhagen V, Denmark. Recuperado de: <http://www.oecd.org/dataoecd/31/13/33684822.pdf> (14 de septiembre de 2011)
- Schultz, T. W. (1961), "Investment in Human Capital", in the *American Economic Review*, Vol. 51, pp. 1-17.
- Secretaría de Educación, Subsecretaría de Planeación. Sistema de Matrícula en Línea.
- Smith, R. (2011). "Multilevel Modeling of Social Problems". Social Structural Research Inc. Cambridge, MA.
- Sosa, Walter (2005). "Perspectivas y avances recientes en regresión por cuantiles". En: *Progresos en econometría*. Temas Grupo Editorial: Asociación Argentina de Economía Política. Págs. 101-138
- Solow, Robert (1956). "A Contribution to the Theory of Economic Growth". *Quarterly Journal of Economics*, 70, pp. 65-94.
- Sund, Krister (2009). "Estimating peer effects in Swedish high school using school, teacher, and student fixed effects". *Economics of Education Review*, 28, pp 329 – 336.

- Tobón, D. et al. (2008). Organización jerárquica y logro escolar en Medellín: un análisis a partir de la función de producción educativa. Lecturas de Economía, Número 68 (enero-junio), pp. 145-17. Recuperado de: <http://aprendeenlinea.udea.edu.co/revistas/index.php/lecturasdeeconomia/article/view/268/215> (15 de septiembre de 2011)
- Toutkoushian, Robert y Curtis, Taylor (2005). "Effects of Socioeconomic Factors on Public High School Outcomes and Rankings". En: The Journal of Educational Research. Vol. 98. No. 5. pp. 259-271. <http://www.jstor.org/stable/27548088> (8 de febrero de 2011).
- Wolff, L., Schiefelbein, E. y Valenzuela, J. (1993). "Mejoramiento de la calidad de la educación primaria en América Latina y el Caribe: hacia el siglo XX". En: Banco Mundial, América Latina y el Caribe, Programa de Estudios Regionales. Informe N° 28.

## Anexos

### Anexo 1. Pruebas para MCO

**Tabla 16. Test de Homocedasticidad matemáticas.**

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: fitted values of matematicas	
chi2(1)	= 194.21
Prob> chi2	= 0.0000

**Tabla 17. Test de Homocedasticidad lenguaje.**

Breusch-Pagan / Cook-Weisberg test for heteroskedasticity	
Ho: Constant variance	
Variables: fitted values of lenguaje	
chi2(1)	= 23.98
Prob> chi2	= 0.0000

Fuente: Cálculo del autor.

**Tabla 18. Varianza Inflada (VIF) matemáticas**

Variable	VIF	1/VIF	Variable	VIF	1/VIF
Nivel madre 4	16,93	0,059067	Ingresos mensuales 2	1,64	0,608895
Nivel madre 3	14,45	0,069193	Nivel madre 9	1,52	0,657756
Nivel madre 2	10,45	0,095681	Nivel padre 5	1,43	0,700312
Nivel madre 1	9,51	0,10511	Nivel padre 7	1,32	0,756885
Nivel padre 4	6,66	0,150093	Ingresos mensuales 4	1,32	0,756999
Nivel padre 3	5,92	0,168878	Nivel padre 9	1,26	0,790525
Nivel madre 6	5,58	0,179108	Tiene computador	1,22	0,817149
Nivel padre 2	4,77	0,209563	Ingresos mensuales 5	1,08	0,927302
Nivel padre 1	4,63	0,215788	Jornada 3	1,07	0,933616
Nivel madre 8	3,94	0,254021	Edad	1,07	0,937018
Promedio lenguaje	2,49	0,402016	Reproductor DVD	1,07	0,938194
Promedio matemáticas	2,44	0,409907	Jornada 4	1,06	0,943967
Nivel padre 6	2,43	0,411645	Género	1,04	0,957998
Nivel padre 8	2,35	0,426128	Ingresos mensuales 6	1,04	0,961091
Nivel madre 5	2,25	0,443511	Personas hogar	1,04	0,963357
Jornada 2	1,96	0,511056	Proviene privado	1,02	0,979389
Jornada 5	1,92	0,519774	Trabaja	1,02	0,984407
Nivel madre 7	1,85	0,540155	Ingresos mensuales 7	1,01	0,988181
Ingresos mensuales 3	1,72	0,581881	<b>Mean VIF</b>	<b>3,34</b>	

Fuente: Cálculo del autor.

**Tabla 19. Varianza Inflada (VIF) lenguaje**

Variable	VIF	1/VIF	Variable	VIF	1/VIF
Nivel madre 4	16,95	0,059003	Ingresos mensuales 2	1,66	0,602124
Nivel madre 3	14,47	0,069113	Nivel madre 9	1,52	0,657429
Nivel madre 2	10,46	0,095596	Nivel padre 5	1,43	0,700095
Nivel madre 1	9,52	0,105053	Nivel padre 7	1,32	0,756773
Nivel padre 4	6,67	0,149879	Ingresos mensuales 4	1,34	0,746891
Nivel padre 3	5,93	0,168664	Nivel padre 9	1,27	0,790399
Nivel madre 6	5,59	0,178896	Tiene computador	2,08	0,481443
Nivel padre 2	4,77	0,209494	Ingresos mensuales 5	1,08	0,924764
Nivel padre 1	4,64	0,215701	Jornada 3	1,07	0,93356
Nivel madre 8	3,94	0,253788	Edad	1,07	0,936475
Promedio lenguaje	2,49	0,40168	Reproductor DVD	1,1	0,91144
Promedio matemáticas	2,44	0,409412	Jornada 4	1,06	0,943694
Nivel padre 6	2,43	0,411097	Género	1,04	0,957728
Nivel padre 8	2,35	0,425528	Ingresos mensuales 6	1,04	0,960726
Nivel madre 5	2,26	0,443273	Personas hogar	1,04	0,963004
Jornada 2	1,96	0,510527	Proviene privado	1,02	0,979321
Jornada 5	1,93	0,519442	Trabaja	1,02	0,984294
Nivel madre 7	1,85	0,53997	Ingresos mensuales 7	1,01	0,987695
Ingresos mensuales 3	1,76	0,569338	Servicio TV	1,17	0,85567
Internet	2,11	0,473212	<b>Mean VIF</b>	<b>3,28</b>	

**Fuente:** Cálculo del autor.

**Tabla20.Test Wald matemáticas**

Test de Wald	
F( 37, 17695)=	66,23
Prob> F =	0,000

**Tabla21.Test Wald lenguaje**

Test de Wald	
F( 39, 17693) =	48,52
Prob> F =	0,0000

**Fuente:** Cálculo del autor

## Anexo 2..

**Tabla 22. Estimación por cuantiles para el área de matemáticas**

	(1)	(2)	(3)	(4)	(5)
VARIABLES	q5	q25	q50	q75	q95
<b>Género</b>	0.0827	2.238***	2.171***	2.103***	2.424***
	(0.986)	(0.524)	(0.369)	(0.297)	(0.258)
<b>Edad</b>	-0.159	-0.584***	-0.610***	-0.566***	-0.628***

	(0.288)	(0.159)	(0.119)	(0.0940)	(0.0978)
<b>Trabaja</b>	-0.0719	-0.177	-0.884	-0.732	-0.923
	(1.994)	(1.139)	(0.924)	(0.754)	(0.615)
<b>Tiene computador</b>	0.299	0.977*	1.001***	1.082***	0.910***
	(1.254)	(0.516)	(0.384)	(0.289)	(0.276)
<b>Reproductor DVD</b>	-0.0733	-0.433	-0.705*	-0.628**	-0.608**
	(0.983)	(0.525)	(0.373)	(0.272)	(0.261)
<b>Ingresos mensuales 2</b>	-0.00713	0.367	0.342	0.329	0.258
	(1.222)	(0.551)	(0.425)	(0.321)	(0.312)
<b>Ingresos mensuales 3</b>	0.0211	0.518	0.839	0.568	0.460
	(1.402)	(0.821)	(0.611)	(0.488)	(0.426)
<b>Ingresos mensuales 4</b>	0.892	1.223	1.500*	1.106*	1.105*
	(2.247)	(1.191)	(0.782)	(0.591)	(0.590)
<b>Ingresos mensuales 5</b>	10.64***	5.839***	3.470***	1.881*	1.297
	(2.596)	(1.861)	(1.236)	(1.052)	(1.152)
<b>Ingresos mensuales 6</b>	15.59***	9.986***	8.826***	7.832***	6.468***
	(2.968)	(2.216)	(1.483)	(1.630)	(1.729)
<b>Ingresos mensuales 7</b>	-14.46	-20.61	-21.55	1.551	0.698
	(13.93)	(14.19)	(13.73)	(13.74)	(12.99)
<b>Nivel padre 1</b>	-0.433	0.306	0.721	0.703	0.634
	(3.634)	(1.574)	(1.103)	(0.846)	(0.793)
<b>Nivel padre 2</b>	0.184	1.662	1.075	0.870	0.607
	(3.753)	(1.461)	(1.065)	(0.813)	(0.781)
<b>Nivel padre 3</b>	-0.104	1.267	0.823	0.754	0.743
	(3.870)	(1.488)	(1.042)	(0.778)	(0.767)
<b>Nivel padre 4</b>	0.00617	1.048	1.198	0.892	0.808
	(3.774)	(1.431)	(0.998)	(0.768)	(0.760)
<b>Nivel padre 5</b>	4.038	2.512	2.956*	2.317*	2.482**
	(7.263)	(2.576)	(1.539)	(1.229)	(1.194)
<b>Nivel padre 6</b>	4.765	2.944*	1.774	2.764***	2.434***
	(4.692)	(1.633)	(1.264)	(0.976)	(0.897)
<b>Nivel padre 7</b>	-0.602	2.966	1.607	1.970	1.549
	(5.152)	(2.721)	(1.722)	(1.625)	(1.440)
<b>Nivel padre 8</b>	4.556	3.786**	2.355**	2.066**	2.314***
	(4.275)	(1.837)	(1.150)	(0.898)	(0.887)
<b>Nivel padre 9</b>	4.835	1.098	3.194	2.643	2.685
	(4.678)	(3.484)	(2.148)	(1.677)	(1.822)
<b>Nivel madre 1</b>	9.242	-1.132	-0.631	0.114	-0.626
	(8.456)	(2.753)	(1.343)	(1.214)	(1.188)
<b>Nivel madre 2</b>	9.123	-1.656	-0.172	0.692	-0.142
	(8.447)	(2.665)	(1.382)	(1.160)	(1.152)
<b>Nivel madre 3</b>	9.374	-1.593	-0.610	0.342	-0.624
	(8.407)	(2.656)	(1.343)	(1.175)	(1.169)
<b>Nivel madre 4</b>	8.881	-1.383	-0.403	0.672	-0.123
	(8.431)	(2.585)	(1.301)	(1.164)	(1.136)
<b>Nivel madre 5</b>	15.24*	-0.0887	-0.0331	0.614	0.206
	(8.419)	(2.766)	(1.488)	(1.398)	(1.456)
<b>Nivel madre 6</b>	10.51	-1.753	1.063	2.054	1.469
	(8.730)	(2.799)	(1.429)	(1.329)	(1.247)
<b>Nivel madre 7</b>	15.36*	2.719	3.652*	4.650***	3.350**
	(8.759)	(3.054)	(1.890)	(1.711)	(1.700)
<b>Nivel madre 8</b>	10.20	-0.725	0.423	1.061	0.128

	(8.560)	(2.872)	(1.490)	(1.309)	(1.246)
<b>Nivel madre 9</b>	15.23*	2.481	2.629	2.578	1.285
	(8.984)	(4.165)	(2.111)	(1.806)	(1.817)
<b>Proviene privado</b>	-1.305	0.932	1.749	1.190**	1.062*
	(3.887)	(1.776)	(1.118)	(0.595)	(0.561)
<b>Personas hogar</b>	-0.0191	0.0360	-0.0180	-0.0446	-0.0348
	(0.222)	(0.134)	(0.0978)	(0.0831)	(0.0775)
<b>Jornada 2</b>	0.255	-0.116	-0.272	-0.236	-0.236
	(1.346)	(0.649)	(0.508)	(0.342)	(0.314)
<b>Jornada 3</b>	-10.49**	-4.618	-3.458	-0.515	-1.065
	(5.295)	(4.347)	(2.957)	(1.728)	(1.286)
<b>Jornada 4</b>	-17.76**	-1.269	-1.088	-1.499	-0.731
	(7.949)	(4.617)	(1.760)	(1.327)	(1.376)
<b>Jornada 5</b>	0.424	0.242	-0.437	-0.277	-0.191
	(1.521)	(0.630)	(0.480)	(0.344)	(0.332)
<b>Promedio matemáticas</b>	0.0279	0.409***	0.263***	0.311***	0.303***
	(0.207)	(0.102)	(0.0819)	(0.0623)	(0.0578)
<b>Promedio lenguaje</b>	0.0474	0.165	0.349***	0.257***	0.285***
	(0.288)	(0.156)	(0.129)	(0.0923)	(0.0878)
<b>Constante</b>	8.349	8.961	10.58**	13.16***	15.17***
	(13.27)	(6.438)	(4.867)	(3.435)	(3.790)
<b>Observaciones</b>	17733	17733	17733	17733	17733
Errores estándar en paréntesis					
*** p<0.01, ** p<0.05, * p<0.1					

Fuente: Cálculo del autor.

### Anexo 3.

**Tabla 23. Estimación por cuantiles para el área de lenguaje**

	(1)	(2)	(3)	(4)	(5)
VARIABLES	q5	q25	q50	q75	q95
<b>Género</b>	0.741	-0.112	0.137	0.182	0.288
	(0.607)	(0.324)	(0.228)	(0.265)	(0.231)
<b>Edad</b>	-0.257	-0.211	-0.223**	-0.224**	-0.274***
	(0.172)	(0.138)	(0.0937)	(0.104)	(0.0692)
<b>Trabaja</b>	0.685	0.206	0.228	0.113	-0.0528
	(0.946)	(1.011)	(0.547)	(0.484)	(0.357)
<b>Tiene computador</b>	-1.682	-0.0983	0.116	0.408*	0.549*
	(1.076)	(0.321)	(0.261)	(0.211)	(0.284)
<b>Reproductor DVD</b>	-0.190	-0.254	-0.0882	-0.0590	-0.126
	(0.771)	(0.307)	(0.227)	(0.213)	(0.258)
<b>Servicio TV</b>	0.0454	-0.157	-0.242	-0.344	-0.188
	(0.645)	(0.410)	(0.293)	(0.232)	(0.255)
<b>Internet</b>	0.655	0.876*	0.585**	0.327	0.486
	(1.163)	(0.465)	(0.286)	(0.321)	(0.335)

<b>Ingresos mensuales 2</b>	3.106***	0.695	0.551	0.265	0.0365
	(0.911)	(0.444)	(0.399)	(0.309)	(0.361)
<b>Ingresos mensuales 3</b>	2.757**	0.722	0.548	0.162	0.0770
	(1.291)	(0.655)	(0.500)	(0.337)	(0.436)
<b>Ingresos mensuales 4</b>	4.230***	1.633	1.335	1.361**	0.888
	(1.279)	(1.040)	(0.848)	(0.689)	(0.683)
<b>Ingresos mensuales 5</b>	5.159***	-1.408	0.626	0.0946	-0.337
	(1.852)	(1.735)	(2.013)	(1.791)	(1.495)
<b>Ingresos mensuales 6</b>	5.158	0.746	3.384	1.924	0.993
	(4.886)	(3.511)	(3.126)	(3.035)	(3.198)
<b>Ingresos mensuales 7</b>	11.61***	6.797***	3.453	5.492***	4.328**
	(2.521)	(1.813)	(2.248)	(1.854)	(1.836)
<b>Nivel padre 1</b>	0.268	-0.0897	-0.644	-0.448	-0.319
	(1.565)	(0.664)	(0.806)	(0.572)	(0.496)
<b>Nivel padre 2</b>	-1.427	-0.550	-0.359	-0.427	-0.212
	(1.516)	(1.038)	(1.011)	(0.634)	(0.561)
<b>Nivel padre 3</b>	-2.422*	-0.472	-0.310	-0.616	-0.385
	(1.396)	(0.705)	(0.785)	(0.580)	(0.468)
<b>Nivel padre 4</b>	-0.604	0.0436	-0.210	-0.00574	-0.0178
	(1.462)	(0.777)	(0.830)	(0.563)	(0.496)
<b>Nivel padre 5</b>	1.792	0.906	0.359	0.885	1.089
	(2.310)	(1.456)	(1.431)	(1.186)	(1.135)
<b>Nivel padre 6</b>	-0.554	1.605	0.973	1.600**	1.052
	(1.625)	(1.184)	(1.013)	(0.705)	(0.725)
<b>Nivel padre 7</b>	-12.63*	1.908*	0.138	0.186	0.571
	(7.505)	(1.028)	(0.889)	(0.930)	(1.088)
<b>Nivel padre 8</b>	2.422	2.407***	1.005	0.562	0.211
	(1.576)	(0.761)	(1.007)	(0.709)	(0.592)
<b>Nivel padre 9</b>	1.765	-1.181	-0.0731	0.538	1.310
	(3.507)	(2.736)	(3.178)	(2.841)	(2.230)
<b>Nivel madre 1</b>	1.598	0.974	0.890	0.809	1.110
	(4.223)	(2.035)	(1.044)	(0.796)	(0.801)
<b>Nivel madre 2</b>	-0.143	0.398	0.546	0.724	0.733
	(4.719)	(1.956)	(1.210)	(0.726)	(0.794)
<b>Nivel madre 3</b>	3.116	0.998	0.896	1.002	1.273*
	(4.372)	(1.953)	(1.029)	(0.702)	(0.773)
<b>Nivel madre 4</b>	2.681	0.891	0.982	1.185	1.291*
	(4.511)	(1.983)	(1.131)	(0.743)	(0.782)
<b>Nivel madre 5</b>	4.608	1.617	2.239**	1.801**	1.852**
	(4.704)	(2.220)	(0.966)	(0.771)	(0.843)
<b>Nivel madre 6</b>	4.576	0.836	1.199	1.398	1.692
	(4.258)	(2.014)	(1.202)	(1.002)	(1.045)
<b>Nivel madre 7</b>	3.782	1.687	3.525***	2.954***	3.350***
	(4.666)	(2.177)	(1.200)	(1.012)	(0.975)
<b>Nivel madre 8</b>	-0.238	0.893	0.693	1.087	1.242
	(5.167)	(2.080)	(1.181)	(0.835)	(0.951)
<b>Nivel madre 9</b>	1.532	-0.660	1.100	0.233	0.253
	(4.855)	(3.728)	(2.287)	(1.472)	(1.439)
<b>Proviene privado</b>	4.135***	2.164***	1.415**	1.280**	0.994**
	(0.887)	(0.509)	(0.586)	(0.596)	(0.490)
<b>Personas hogar</b>	-0.302	-0.215**	-0.111*	-0.0541	-0.0652
	(0.206)	(0.0897)	(0.0570)	(0.0526)	(0.0534)

<b>Jornada 2</b>	0.527	0.461	0.287	0.394	0.478*
	(1.043)	(0.398)	(0.374)	(0.264)	(0.260)
<b>Jornada 3</b>	-1.130	-3.668**	-2.310*	-2.830***	-2.793***
	(1.907)	(1.453)	(1.322)	(0.927)	(0.852)
<b>Jornada 4</b>	-5.660*	-1.083	0.436	-0.398	-0.756
	(3.428)	(2.869)	(1.413)	(0.817)	(0.713)
<b>Jornada 5</b>	1.295	-0.290	-0.117	-0.173	-0.155
	(1.101)	(0.384)	(0.300)	(0.213)	(0.259)
<b>Promedio matemáticas</b>	0.0522	0.131*	0.175***	0.170***	0.149***
	(0.161)	(0.0784)	(0.0517)	(0.0536)	(0.0540)
<b>Promedio lenguaje</b>	0.372**	0.407***	0.275***	0.277***	0.315***
	(0.157)	(0.144)	(0.0787)	(0.0738)	(0.0794)
<b>Constante</b>	10.73	11.93**	18.51***	19.79***	20.33***
	(11.84)	(4.724)	(3.221)	(2.990)	(2.972)
<b>Observaciones</b>	17733	17733	17733	17733	17733
Errores estándar en paréntesis					
*** p<0.01, ** p<0.05, * p<0.1					

**Fuente:** Cálculo del autor.